

Contra Costa County's Transitioning Out to Stay
Out (TOSO) Program for Justice-Involved Youth

MENTALLY ILL OFFENDER CRIME
REDUCTION (MIOCR) GRANT:
FINAL REPORT

Prepared by Kate Shade, PhD and Damir Ceric, MS

With assistance from Eva Gonzales, Sara Lapicola, and Deni Yandell

Special thanks to James Rivers

September 2018

EXECUTIVE SUMMARY

This is the final evaluation report for a three-year project titled ‘Transitioning Out to Stay Out’ (TOSO) program. The core of the TOSO intervention was delivery of Functional Family Therapy (FFT) for juvenile probationers returning to the community after out-of-home placement in Contra Costa County. Process evaluation measures included regular review of referral and enrollment services and an assessment of fidelity to the FFT model. Summative evaluation measures included pre-post examination of change among FFT youth and family members. The primary outcome was rearrest among TOSO youth who were compared to youth who received Multisystemic Family Therapy (MST) and sustained petitions for completers versus non-completers who received FFT. The cost of FFT and MST was also determined to inform other counties where FFT and/or MST may be provided to justice-involved youth.

The TOSO project was envisioned in response to the California Board of State and Community Correction’s (BSCC) Mentally Ill Offender Crime Reduction (MIOCR) grant request for proposals. Primary program partners included Contra Costa County’s Behavioral Health Children’s System of Care and the Probation Department as well as a community-based organization, Community Options for Families and Youth (COFY).

The TOSO program goals were to transition youth from out-of-home placement back to their families and their communities, providing FFT to increase protective factors and decrease risk factors related to recidivism.

The target population for TOSO was youth at highest risk for recidivism. The population has been remanded to in-custody treatment due to chronic offending, exposure to high risk sex work and exploitation as a minor, and/or as commission of a violent crime.

TOSO program youth were expected to show the following improvements after participation: 1) reduced criminal activity among Contra Costa County in-custody treatment graduates, 2) a reduced the number of girls who exit treatment and are re-arrested for prostitution-related offenses, 3) a reduced number of out of home non-County placements for youth through age 19 years, 4) improved high-risk youth and family functioning, and 5) an increased number of high-risk youthful offenders who attend school and/or employment.

A process evaluation indicated that effectiveness was improved through strong connections between the probation officers, behavioral health liaisons, and FFT clinicians implementing TOSO. Clear referral processes, contact with youth and families while the youth were still in-custody, and

consistent messaging about the purpose of FFT for court-mandated youth improved engagement and treatment completion. Developing clinical competence and fidelity to the model of FFT was challenged by turnover among the FFT clinicians and supervisor; consistent with research in the field of implementation science, the team stabilized, clinical competence was established, and adherence to FFT best practices were improved by year two of the TOSO project.

An impact evaluation demonstrated that MST was more effective than FFT in reducing recidivism among probationers. The result of a logistic regression to assess the effect of the two programs demonstrated that FFT youth were 1.3 times (95% confidence interval of 0.7-2.345) more likely to be rearrested than MST youth. Greater time in treatment also predicted a reduction in rearrest. For the 57 youth who completed MST and the 55 youth who completed FFT and were then arrested 9, 18, or 24 months after treatment, the number needed to treat with MST versus FFT to reduce one rearrest ranged from 2-13 youth.

In evaluating only the 272 youth who participated in FFT, we found that 22 youth had sustained petitions following referral to and exit from TOSO; the 22 were equally distributed between those who did (13 youth, 10%) and those who did not (9 youth, 10%) complete FFT. The majority were low-income (MediCal eligible) males from two impoverished areas of the county, of moderate risk, and referred after being remanded to the Orin Allen Rehabilitation Facility, commonly known as the ranch.

TOSO achieved most of the targets established when the MIOCR-funded project was conceived: 1) recidivism was reduced among youth exiting in-custody treatment programs, 2) only one youth was returned to the Girls in Motion (GIM) treatment program and none to the Youthful Offender Treatment Program (YOTP), 3) TOSO is on track to graduate 60% of families referred for FFT, 4) high and moderate risk youth showed significant improvements in behavior, mental health symptoms, and family functioning after completion of FFT, and 5) at exit from TOSO, 72% of youth were either attending school, employed, or both.

Barriers were primarily related to the absence of an effective tracking mechanism for youth probationers. It is difficult to collect, manage, store, share, evaluate and respond to trends in criminal activity in a timely way without an effective case management system for juvenile probation. There were several limitations to our analyses and the results should be interpreted with caution. We had no consistent measure of recidivism and limited data about pre-TOSO desistance from crime among the youth returning from in-custody programs, especially the ranch. Most importantly, we had inadequate measures of risk among the MST and FFT samples and suspect that there were covariate factors we could not account for that might explain the

difference in rates of rearrest. Though we had adequate sample sizes, the outcome data were not normally distributed. This is good news, in that it indicated low rates of rearrest or sustained petitions in the two groups of youth. However, our analyses were also limited due to the varied and relatively short period for follow-up.

The primary lesson learned was how critical strong and effective partnerships are in planning, implementing, evaluating, and improving a program such as TOSO. All teams involved in the project focused on timely, consistent, and effective communication between probation officers, the juvenile court, in-custody treatment probation and mental health staff, community-based mental health liaisons, the FFT clinicians, and, most importantly, the youth and families.

As a result of the impact evaluation, we recommend that the probation department utilize a reliable and valid risk assessment tool, such as the Ohio Youth Assessment System's (OYAS) re-entry risk assessment for case planning in preparing youth for their return to the community. Since the majority of referrals and reoffenders were exiting the ranch, improvement is most critically needed in assessing risk, needs, and responsivity for these youth. Youth in this sample self-reported serious mental health problems such as self-harm, severe anxiety, and psychosis. In order to provide early and effective treatment, an additional tool should be used to identify serious psychiatric symptoms when youth are placed in-custody. The Structured Assessment for Violence Risk in Youth (SAVRY) has been validated with serious youthful offender populations. We recommend that the probation department also increase the dosage of evidence-based cognitive-behavioral treatments at the ranch, ensuring that every youth participates in all sessions of Aggression Replacement Training and Thinking for a Change prior to release.

As a result of the process evaluation, we suggest that ongoing, supportive partnerships are essential to sustaining the delivery of family therapy to justice-involved youth. The mental health liaisons were critical to the implementation of TOSO as designed, enabling youth and families to engage with FFT clinicians while still in-custody. The partnerships developed between probation, mental health, and the Community Options for Families and Youth (COFY) staff should be further strengthened through joint trainings, regular case staffing, and shared professional self-care activities. It falls upon the supervisors of the front-line probation, mental health, and COFY personnel to ensure that staff are well-trained, achieve competence in working with youth probationers, feel confident about the quality of their work, are recognized for their achievements, and are effectively managing the stress associated with this work. Peer and supervisor support are essential to the prevention of compassion fatigue and burnout.

TABLE OF CONTENTS

INTRODUCTION

History of Family Therapy for Juvenile Offenders in Contra Costa County	10
Data Supporting the Need for the Program	11
Purpose and Objectives of the Program	11
County-Community Partnerships	12

REVIEW OF THE LITERATURE

Multisystemic Family Therapy (MST)	13
Functional Family Therapy (FFT)	14
Economic Evaluations of MST and FFT	14

RESEARCH DESIGN

Youth and Families	15
Criteria for and Process of Referral to Program	17
Design of Process Evaluation	18
Design of Impact Evaluation	18

ANSWERING THE RESEARCH QUESTIONS

Research Questions	19
Independent Variables	19
Conceptual and Operational Definitions	19
Dependent Variables	20
Conceptual and Operational Definitions	20

DATA COLLECTION AND ANALYSIS

Data Sources	21
Data Collection Methods	22
Measurement Tools	22
Data Limitations	32
Data Analysis for Impact Evaluation	33

PROCESS EVALUATION RESULTS	
Coordination of Services	34
Timeliness of Treatment	34
Fidelity to the FFT Model	35
IMPACT EVALUATION RESULTS	
MST and FFT Probation Youth	39
Final MST and FFT Sample	40
MST versus FFT Rearrests	44
MST versus FFT Rearrests (2016-2018)	46
FFT Participants	49
FFT Treatment Teams	51
FFT Youth Prior Involvement with Juvenile Justice System	52
FFT Youth Completers versus Non-Completers	53
Youth and Family Functioning Before and after Completion of FFT	56
FFT Youth Recidivism	64
TOSO Program Success	69
Costs of TOSO	70
CONCLUSIONS AND RECOMMENDATIONS	71
REFERENCES	75
TABLES	
Table 1. Youth Targeted for TOSO	14
Table 2. Clinical Competence and FFT Adherence according to TYPE Reports	37
Table 3. Mean Scores for the Client Outcome Measure-Parent Report (COM-P)	38
Table 4. Mean Scores for the Client Outcome Measure-Adolescent (COM-A)	39
Table 5. Gender for MST versus FFT Youth	41
Table 6. Age for MST versus FFT Youth	41
Table 7. Mental Health Disorders for MST versus FFT	42
Table 8. JAIS Categories for MST versus FFT Youth	42

Table 9. CALOCUS/LOCUS Level of Need (Category) for MST versus FFT Youth	43
Table 10. Gender, JAIS, and Mental Health Disorder, for Completers (2016-2018)	47
Table 11. Age of MST and FFT Completers (2016-2018)	47
Table 12. Number of Charges for first Rearrest for MST and FFT Completers (2016-2018)	48
Table 13. Charges by Crime Type for first Rearrest for MST and FFT Completers	48
Table 14. Relative Risk Reduction, Absolute Risk Reduction and Number Needed to Treat	48
Table 15. Age of TOSO Youth	49
Table 16. Gender of TOSO Youth	50
Table 17. Racial/Ethnic Group of TOSO Youth	50
Table 18. Area of County of TOSO Youth	50
Table 19. Number referred to TOSO, enrolled, completed (% completed after opening), by Treatment Team	51
Table 20. Age of First Sustained Petition and First Detention for TOSO Referrals	52
Table 21. Reasons for Non-Completion of FFT	53
Table 22. Referral Source, by FFT Completion Status	54
Table 23. JAIS Risk Category, by FFT Completion Status	54
Table 24. Age of Youth, by FFT Completion Status	55
Table 25. Racial/Ethnic Group, by FFT Completion Status	55
Table 26. Correlations between Prior Justice System Involvement and FFT Completion	56
Table 27. Pre-FFT Outcome Questionnaire Subscale and Total Scores	57
Table 28. Pre-FFT Youth Outcome Questionnaire Subscale and Total Scores	58
Table 29. Pre-FFT Youth Outcome Questionnaire-Self Report Subscale and Total Scores	58
Table 30. Pearson Product-Moment Correlations of Pre-FFT YOQ and YOQ-SR	59
Table 31. Post-FFT Outcome Questionnaire Subscale and Total Scores	60
Table 32. Post-FFT Youth Outcome Questionnaire Subscale and Total Scores	60
Table 33. Post-FFT Youth Outcome Questionnaire-Self Report Subscale and Total Scores	61
Table 34. Sustained Petitions for Property Offenses 6-18 months after Exit from TOSO	65
Table 35. Pre-Post Sustained Petitions, by FFT Completion	66
Table 36. Number of TOSO Youth with a Sustained Petition for a Serious Charge, by Completion Status	68
Table 37. Cost to Deliver MST and FFT to Probation Youth in Contra Costa County	70

FIGURES

Figure 1. Mean CALOCUS/LOCUS Total for MST versus FFT Youth	44
Figure 2. Kaplan-Meier Survival Curves for Days to Rearrest for FFT and MST Youth	46
Figure 3. Box Plot of ranked values for Pre-Post YOQ-SR Scores for the Critical Items	62

REFERENCES	75
------------	----

APPENDICES	90
------------	----

Appendix A. Logic Model for Contra Costa County's MIOCR-funded TOSO Project	
Appendix B. Child and Adolescent Level of Care Utilization System (CALOCUS) Worksheet	
Appendix C. Categories for Level of Care based on Composite CALOCUS/LOCUS Score	
Appendix D. Contra Costa County Residence Zip Codes for TOSO Youth Referrals and those with Sustained Petitions	

Final Contra Costa County Program Report
Mentally Ill Offender Crime Reduction (MIOCR) Grant Program
Transitioning Out to Stay Out (TOSO)

INTRODUCTION

Functional Family Therapy (FFT) has long been established as an evidence-based treatment modality for juvenile offenders with conduct and/or substance use disorders (Henggeler & Sheidow, 2012). The treatment is provided in three distinct clinical phases: Engagement and Motivation, Behavior Change, and Generalization. During each phase different assessment strategies, therapy goals, and therapist skills are utilized to guide families through a systematic process of change. The goals of the first phase are to engage the family members in a therapeutic alliance, reduce shaming and blaming, and establish an optimistic outlook about making changes. Validating feelings and reframing behaviors and motivations are pertinent therapeutic skills used during this phase. The second phase focuses on identifying and targeting specific risk behaviors. The therapist teaches and reinforces effective family functioning, focusing on skills such as clear communication, parental supervision and parental support, conflict management, and problem solving. The goals of the second phase are to reduce risk and increase protective factors. Specifically, the clinician works with the family to decrease neglectful, harsh and coercive parenting and to improve the consistency and productivity of parent-child communication. During the third and final phase, the therapist and family members work to generalize the positive behaviors practiced within the family to relations outside of the family. Interactions are addressed with extended family, school personnel, the juvenile justice system, and relevant support persons in the broader community. FFT is a model program (Blueprints for Healthy Youth Development, 2015) recognized for its effectiveness in reducing recidivism, even among the most serious youthful offenders.

ACRONYMS: Cognitive Behavioral Therapy (CBT); Commercially Sexually Exploited Youth (CSEY); Community Options for Families and Youth, Inc. (COFY); Functional Family Therapy (FFT); Girls in Motion (GIM – in-custody treatment program that targets high-risk females); Multisystemic Family Therapy (MST); Orin Allen Youth Rehabilitation Facility (OAYRF – also known as ‘the ranch,’ in-custody treatment program for moderate-risk males); Youth Offender Treatment Program (YOTP – in-custody treatment program that targets high-risk males)

History of Family Therapy for Juvenile Offenders in Contra Costa County

Since 2013, Multisystemic Family Therapy (MST) has been offered to Contra Costa County families with 12-17 year olds who are at a moderate to high risk of offending. Multidimensional Family Therapy (MDFT) is provided to 12-17 year old juvenile offenders with substance abuse. MST and MDFT are funded through the Mental Health Services Act, (MHSA) developed in response to voter approval of Proposition 63 in November of 2004 (Contra Costa County Health Services, 2018). A community-based organization, Community Options for Families and Youth (COFY) contracted with the County to provide MST to youth probationers and those at risk of involvement with probation, for youth aged 12-17 years. Prior to the implementation of the Transitioning Out to Stay Out project (TOSO) , there were no family therapy services available to families with high risk 18- to 20-year old male offenders exiting in-custody treatment or non-County placements; there were also no family therapy services for 12-20 year old commercially sexually exploited youth (CSEY). The Mentally Ill Offender Crime Reduction (MIOCR) grant for juveniles, awarded to the Contra Costa County Probation Department in 2015, launched a sustainable family therapy project designed to improve transitions for youth returning from custody to the community, and reduce recidivism among high-risk youth and families.

Data Supporting the Need for the Program

Functional Family Therapy (FFT) complements the in-custody treatment of youth probationers and increases treatment intensity for the County's highest risk juvenile offenders. Successful transitions from the County Probation Department's juvenile treatment programs (YOTP, GIM, and the ranch), and non-County placements are required to decrease the total serious crime committed by juveniles in the County. FFT is designed to provide services that reduce risk factors which result in out of home placement and that increase protective factors which support youth in desisting from crime. State-wide data available at the time of our submission of the MIOCR grant application indicated that, in 2013, 37,615 youth in the entire state of California became wards of the court and 23% were detained (Harris, 2014). In 2015, wardship decreased; 28,447 youth became wards of the court. However, the percent detained in the state of California increased from 23% to 30.2% (Harris, 2016). In 2017, wardship again decreased to 23,689 but, once again, the percent detained remained higher (29.9%) than the percent in 2013 (Becerra, 2018).

In our application for MIOCR funding (Contra Costa County Probation, 2015) we also cited the most recent juvenile crime statistics for the County. The violent crime rate for juveniles in Contra Costa County in 2013 was 37% for all felony arrests, ten percent greater than for the entire state of California (kidsdata.org, n.d.a). For the year of last report, the incidence of violent crime by juveniles was 39%, still higher than the incidence rate (34%) for the entire state (kidsdata.org, n.d.b). According to the Center on Juvenile and Criminal Justice's (2018a) analysis of California counties' response to juvenile crime in 2016, Contra Costa County had an average daily population in its juvenile hall and ranch of 148 youth as well as 19 transfers to the Department of Juvenile Justice (DJJ). Only Tuolumne County transferred more juveniles to DJJ than Contra Costa County. Contra Costa made 106.6 youth transfers per 1,000 juvenile felony arrests at an estimated cost of \$271,675/ward (Office of the Governor, 2017). The average rate of transfer for all counties in California was 34.3/1,000 juvenile felony arrests (Center on Juvenile and Criminal Justice, 2018a). During 2016-2017, Contra Costa County committed more juveniles to DJJ than Alameda, Marin, San Francisco, and San Mateo combined (Center on Juvenile and Criminal Justice, 2018b). These data suggest that projects such as TOSO are needed to reduce violent crime, decrease detentions among probation wards, improve transitions from detention to the community, save taxpayer dollars, and reduce recidivism.

Purpose and Objectives of the Program

The target population for TOSO is youth at highest risk for recidivism. The population has been remanded to in-custody treatment due to chronic offending, exposure to high risk sex work and exploitation as a minor, and/or as commission of a violent crime. The primary purpose of TOSO is to transition youth from out-of-home placement back to their families and their communities, providing family therapy to increase protective factors and decrease risk factors related to recidivism. The specific objectives are 1) reduce recidivism among Contra Costa County in-custody treatment (YOTP, GIM and ranch) graduates, 2) reduce the number of girls who exit treatment and are re-arrested for prostitution-related offenses, 3) reduce out of home non-County placements for youth through age 19, 4) improve high-risk youth and family functioning, and 5) increase the number of high-risk youthful offenders who attend school and/or employment (see the Logic Model for MIOCR Project, Contra Costa County, Appendix A).

When establishing the TOSO program, we set the following targets for year four: 1) 20% reductions in re-arrests among YOTP, GIM and ranch participants over the years 2015-2018, 2) not more than 10% of GIM participants returning to the GIM program by 2018, 3) completion of FFT by more than 70% of families referred, 4) maintenance of the youth in the home with significant improvements in pre-post YOQs for those completing FFT and 4) 75% of youth in families participating in FFT attending school and/or employment at completion of program. The overall goal of the project will have been achieved if the benefits of desistance from crime outweigh the costs of the program.

County-Community Partnerships

Contra Costa County developed a comprehensive plan for providing a cost-effective continuum of services as directed by the Mental Health Services Act (MHSA); the plan was amended in 2012 and again in 2018 (Contra Costa County Health Services, 2018). The continuum includes primary, secondary and tertiary prevention services for children and adolescents with behavioral disorders. It was anticipated that the expansion of services through MHSA would reduce residential placements, suicide and self-harm, school failure and early exit from school, referrals to child welfare services, homelessness, and unemployment for County youth. In planning for the 2018 amendment, staff from Contra Costa County's Behavioral Health Services evaluated the county-wide service delivery processes and youth outcomes.

The County uses a Child and Adolescent Level of Care Utilization System (CALOCUS/LOCUS) to assess individual's needs when seeking behavioral health services (Contra Costa County Health Services, 2018). In evaluating MHSA spending for fiscal year 2015-2016, the County determined that 1) the number of youth served met projections set in 2012, 2) the County was serving more youth in need of behavioral health services than the majority of California counties, and 3) most MHSA funds were being spent on tertiary prevention services rather than lower level, primary and secondary prevention services for youth. Prevention services include programs such as MST and FFT; high intensity, community-based interventions designed to support children and adolescents in the home. The TOSO project, a partnership between Contra Costa County's probation department, behavioral health children's system of care, the juvenile court, and a community-based service provider, COFY, has enhanced county-community partnerships that benefit youth and families.

REVIEW OF THE LITERATURE

Multisystemic Family Therapy (MST)

Multisystemic Family Therapy (MST) is an intensive home-based family therapy intervention designed to reduce risk factors, improve protective factors and address the needs of youth offenders in the context of individual-, family-, peer-, school-, and community-level systems (Henggeler, 1997). Developed by Scott Henggeler and associates, MST therapists provide a range of goal-directed services to each family. The aim of the intensive program is to improve parenting practices, increase expression of affection within the family, decrease youths' contact with antisocial peers, increase association with prosocial peers, improve attendance and performance in school, encourage involvement of the youth and family in positive social and recreational activities, teach healthy communication techniques, and support parents in monitoring youth and problem-solving on behalf of the needs of the family. Treatment is provided by trained clinicians with small caseloads who are available 24 hours a day, seven days a week to offer crisis intervention services. Eventually, as crises lesson and functioning improves within the multiple systems in which the family is located, services are terminated; usually this is accomplished within 4-6 months (Henggeler & Lee, 2003).

A significant body of literature supports the impact of MST on reducing recidivism among youthful offenders. In an early study in which Henggeler and colleagues were directly involved in implementation and evaluation, recidivism was reduced by 43% and out-of-home placements by 64% for 84 youth with chronic and violent offenses (Henggeler, Melton, & Smith, 1982). More recently, results of a randomized, clinical trial conducted independent of the MST developers showed a 38% decrease in rearrests among 93 probationers during 18 months of follow-up (Timmons-Mitchell, Bender, Kishna, & Mitchell, 2006). In another randomized, clinical trial with justice system-involved youth there was a 55% difference in rearrests and a 57% difference in detentions over a 14 year period of follow-up for youth who received MST versus individual therapy (Schaeffer & Borduin, 2005). Importantly, Schoenwald and Henggeler (2004) demonstrated criminal offenses among juvenile offenders could be reduced by 35% when therapists adhered to the MST model and 53% when supervisors implemented MST with fidelity to the model.

Functional Family Therapy (FFT)

Functional Family Therapy (FFT) is a short-term home-based treatment that engages caregivers, increases motivation to change, improves parental monitoring, reduces family conflict, and establishes connections that support the family within the community (Sexton & Turner, 2011). The therapy integrates a cognitive-behavioral approach while maintaining a focus on improving functional relationships in the family. FFT has strong evidence of positive outcomes including reductions in substance use, out of home placements, and recidivism for the target child as well as other children in the home (Klein, Alexander & Parsons, 1977; Waldron, Slesnick, Brody, Turner & Peterson, 2001).

FFT is an effective treatment for juvenile offenders. A large controlled trial of youth remanded to probation and randomized to FFT or supervision as usual found a 30% reduction in violent crime among those who received FFT as compared to controls. However, this effect was only true when FFT was delivered by a competent therapist who adhered to the treatment guidelines (Sexton & Turner, 2011). Another randomized controlled trial tested the effectiveness of FFT for both girls and boys and found reductions in violent criminal behavior well into adulthood (Gordon, Graves, & Arbuthnot, 1995).

Many studies have established the role that poor parent-child relationships play in the development of delinquency (Cook & Gordon, 2012; Dodge, Coie, & Lynam, 2006). Households in which there is low supervision and absent or inconsistent parental monitoring of children's activities facilitates participation in a deviant peer group, enables substance use, and is associated with poor school performance. These are risk factors for both girls and boys involved in the juvenile justice system (Masseti et al, 2011; Thornberry, Freeman-Gallant, & Lovegrove, 2009). FFT is designed to address these familial risk factors and promote positive and supportive parent-child relationships.

Economic Evaluations of MST and FFT

A number of studies have been conducted to examine the costs and benefits of MST and FFT. Additional studies have compared the two. The most extensive reports have been published by the Washington State Institute for Public Policy; researchers have estimated that the cost of delivering

MST or FFT to families in Washington, based on 2003 dollars, to be \$5,681 for MST and \$2,140 for FFT (Aos, Lieb, Mayfield, Miller, & Pennucci, 2004). Costs for the two programs differ due to the suggested caseloads full-time clinicians can carry. In the case of MST, since clinicians are expected to be available 24/7 for crisis intervention, the MST team consists of 2-4 therapists with a supervisor, with each clinician working with no more than 6 families at a time (MST Services, 2009). Aos et al. (2004) also calculated the benefits, minus the costs, for a family participating in MST or FFT, finding a benefit/cost estimate of \$9,316 for MST and \$14,315 for FFT. Goorden et al. (2016) conducted a systematic review of 11 cost evaluation studies of family therapy with youth at risk for and/or involved with the justice system, including one of FFT and eight of MST. The researchers found most of the studies reviewed were of poor to fair quality, as measured using the British Medical Journal Checklist (Drummond & Jefferson, 1996) and the Consensus on Health Economic Criteria checklists (CHEC; Evers, Goossens, de Vet, van Tulder, & Ament, 2005). Goorden et al. (2016) reported that no researchers accounted for all relevant costs and benefits in their analyses.

Recently, Baglivio, Jackowski, Greenwald, and Wolff (2014) evaluated the effectiveness of MST versus FFT for justice-involved youth in Florida, using a propensity score matched pair design to ensure the two groups were comparable in terms of risk for re-offending. Baglivio et al. (2014) found no significant difference in rearrests, reconvictions, or violations of probation among the MST versus FFT samples. Results of this study suggest that MST and FFT may be similarly effective and, by virtue of its lower cost, FFT is the more cost effective treatment.

RESEARCH DESIGN

Youth and Families. TOSO was designed to provide transitional services to the serious youth offenders from YOTP and GIM as they returned to their families following in-custody treatment (see Table 1). The mainstay of the current treatment provided to the youth in YOTP and GIM is based on a cognitive-behavioral treatment (CBT) approach. Researchers have demonstrated that CBT, regardless of the brand, effectively reduces recidivism by 25% (Landenberger & Lipsey, 2005). Reductions upward of 40% are possible when the dose is adequate and the treatment is implemented with adherence to the model. When youth are sent home from court-mandated out-of-home placements, their family and community environments are often little changed. Enhanced and extended cognitive-behavioral treatment is effective at reducing recidivism among youth at high risk

for a return to custody and/or entry into the justice system as an adult (Landenberger & Lipsey, 2005).

TABLE 1
Youth Targeted for TOSO

	Age	Race	Area of County	Prior Placements	Age of 1st Arrest	Total Served
YOTP Program (males) launched 2008	15-20yrs average 16.8	57% Black 23% Latino 10% White 10% Other	Majority from East County (Pittsburg, Antioch, Bay Point)	Average 5 prior juvenile hall detentions, 2 out of home placements	6-19 years, average 14	156 w/12 non-completers
GIM Program (females) launched 2010	13-23yrs average 16.7	55% Black 23% Latino 16% White 6% Other	50% from East County (Pittsburg, Antioch, Bay Point), 48% from West County (Richmond)	0-12 prior juvenile hall detentions, average 3	11-18 years, average 15	160; 38 (24%) with 2 or more GIM commitments

Note: YOTP evaluation completed for youth remanded to program 2008-2013; GIM evaluation completed for youth remanded to program 2010-2016. Ranch youth were also targeted for TOSO.

In 2013, the County increased the CBT dosage for youth in YOTP and GIM and implemented oversight and quality improvement processes to improve the delivery of evidenced-based practices. For the 90 juvenile offenders who completed YOTP prior to the increase in dosage of CBT (2009-2012), 68% recidivated. Among the 66 participants who completed YOTP since 2013, 62% recidivated. Forty-two percent of the group who exited the program during the first three years was re-arrested for a violent offense while 35% among the group who exited during the last three years had a violent charge upon re-arrest (Gerchow, 2015).

We conducted a recent evaluation of 160 GIM participants who were remanded to the program between January 2010 and May 2016 (Shade & Ceric, 2017). The period of time for follow-up to measure recidivism varied due to the different exit dates for participants. Unfortunately, some youth were rearrested within 1-2 months after release from GIM. For youth who were not rearrested, their time in the community ranged from 1.8-75 months (0.2-6.3 years) of follow-up. Of the 160 GIM participants who exited the program before May of 2016, 77 (48%) were rearrested and 83 (52%) were not. We also determined the average time-to-rearrest for only those youth who had been out in the community following release from GIM for at least 6 months. We found that commercially sexually exploited youth (CSEY) had an almost two-fold increase in the odds of rearrest than non-

CSEY (odds ratio 1.97; 95% confidence interval of 0.96 – 4.04, n.s.) By approximately 200 days after return to the community, GIM youth with CSEY involvement recidivated; of the 138 GIM youth rearrested between July 2010 and January 2016, 70% had been rearrested by day 250. These data suggest enhanced probation supervision and increased family support are needed during the critical 6 month transition from in-custody to the community to reduce recidivism in this high risk population.

Criteria for and Process of Referral to TOSO. Youth were referred to TOSO if they were 1) exiting YOTP, 2) exiting GIM, 3) former YOTP youth or 4) former GIM youth who remained on probation, 5) siblings of YOTP or GIM youth who were on probation, 6) youth returning from non-County placement, 7) youth released from the ranch, and 8) CSEY who were working with a case management service provider in the community.

An essential component of the success of the TOSO project was the established relationship between the staff employed by the County's behavioral health children's system of care and juvenile probation departments. After an early false start, during which youth were referred to TOSO who were not the target population for MIOCR services, the FFT clinical, mental health, and probation teams clearly communicated the aim of TOSO to service providers throughout the county. FFT clinicians were formally introduced to in-custody mental health and probation staff and maintained regular communication in order to facilitate timely referrals. Central to the TOSO model was outreach and engagement of youth and family while the youth was still in custody. Three liaisons for the behavioral health children's system of care, stationed at each of the three field probation offices (in East, Central, and West Contra Costa County) assisted with the referral process.

Typically, FFT is delivered in 12-15 sessions over 8-26 weeks, though outreach activities prior to the first session may require several contacts. The liaisons were in regular communication with probation and mental health staff serving the in-custody youth, the youth and parent(s), and also with the FFT clinicians, to facilitate early engagement of youth/families. They assessed whether family therapy was feasible, identified other services that were needed or were already being provided to the youth/family, obtained information about MediCal status, determined risk level, anticipated the date of release for youth in-custody, communicated about any transfer of services between probation officers, assisted with transportation, and asked about a preferred contact,

language, meeting location, and meeting time for the family. They explained the objectives of FFT to the family members and provided a warm hand-off to the FFT clinicians. Liaisons worked with all involved to ensure 1-2 sessions were scheduled for youth while they were still in custody.

Design of Process Evaluation

The process evaluation consisted of a mixed methods design. We included fidelity measures and MIOCR grant-funded targets and timelines to assess the strengths and weaknesses of TOSO during its implementation. Program activities were implemented, evaluated and re-structured (based on the process evaluation data) by the project manager, evaluator, FFT clinical team, and the national office of FFT in consultation with the County mental health staff and with formative data and analysis supplied by associates from the California Institute for Behavioral Health Solutions. Process evaluation data retrieved and analyzed by the FFTI clinical supervisor and shared with the project manager and evaluator included entry rate, non-completion rate and reason, utilization rate, waitlist information, number of youth exiting the program, and rate of successful completion of FFT. The project manager also tracked information about billing, the percent of families enrolled in TOSO who were MediCal-eligible, and reimbursement to the County for services provided to youth with MediCal. The process evaluation was facilitated by monthly case review meetings which all team members attended. At this meeting, problems such as difficult communication between youths' probation officers and FFT clinicians or limited referrals from an individual source were discussed and addressed.

Design of Impact Evaluation

The impact evaluation was conducted using a quasi-experimental design. We used data collected from two samples to 1) compare the effectiveness of MST versus FFT, with any rearrest following treatment as the outcome of interest and 2) examine the effect of FFT on the primary outcome, recidivism, measured as a sustained petition following treatment. For the FFT-only analysis, we examined the effect of FFT completion, using a pre-post design in which participants served as their own controls on the primary outcome, number of sustained petitions, and several secondary outcomes associated with recidivism. These included parent(s) and youth behavior, mental health symptoms, and intra- interpersonal functioning (Barnoski, 2004).

RESEARCH QUESTIONS

We aimed to answer the following questions in evaluating the impact of the TOSO project.

1. How effective was FFT, compared to MST, in reducing recidivism, measured as rearrest for any charge and for violent, property, and weapon charges at 18 months?
2. How effective was FFT at reducing recidivism, measured as a sustained petition and/or a return to custody for completers versus non-completers of FFT?
3. How effective was FFT at reducing violent, property, and weapon crimes for TOSO youth who successfully completed FFT?
4. Was there a dose-response relationship, with length of time in treatment instead of treatment completion predicting any post-FFT sustained petitions?
5. How effective was FFT in improving youth behavior, psychological symptoms, and family functioning?
6. How effective was FFT at increasing youth's attendance in school and employment?

Independent Variables. The main independent variable of interest was participation in family therapy. A number of covariates were included, based upon research about the predictors of recidivism among youthful offenders. These included age, gender, racial/ethnic group, neighborhood, prior involvement in the justice system, school attendance, and employment (Murray & Farrington, 2010).

Conceptual and Operational Definitions. The main independent variable was conceptualized as participation in a family therapy intervention known to reduce recidivism among youth offenders (Alexander & Sexton, 2000; Henggeler, Milton, & Smith, 1992). The variable was operationalized as completion of MST or FFT, as designed. For FFT participants, we collected information about dose of treatment as well as completion of treatment; completion was defined as at least 60 days in treatment and the achievement of the three phases of FFT. Prior involvement in the justice system was conceptualized as the depth and breathe of exposure to the juvenile justice system for TOSO youth. The variable was operationalized as 1) age of first sustained petition, 2) age of first detention, 3) source of referral (with ranch, YOTP and GIM youth likely to have a more extensive history of involvement) and 4) number of prior detentions or court-ordered out-of-home placements.

Covariate risk for recidivism was conceptualized as additional variables associated with criminal behavior and involvement in the justice system among youth. They were operationalized as 1) age at time of referral; 2) gender, as reported at time of referral; 3) racial group, as reported at time of referral; 4) area of county, determined by residential zip code at time of referral; 5) enrollment and attendance in school at time of referral; and 6) employment at time of referral.

Dependent Variables. The main dependent variable was recidivism. Data were also collected to explore secondary outcomes among FFT youth related to parent and youth functioning within the family.

Conceptual and Operational Definitions. Recidivism was conceptualized as a return to criminal behavior following participation in TOSO, an intervention designed to reduce participation in crime (National Institute of Justice, 2014). Recidivism can be operationalized in a number of ways; it is best assessed using multiple measures (Council of State Governments Justice Center, 2014; van Batenburg-Eddes et al., 2012). For the MST versus FFT analysis, recidivism was measured using the Criminal Offender Record Information (CORI) section of the California Law Enforcement Telecommunications System (CLETS), provided through the probation department. Official arrest records are an acceptable measure of recidivism, especially for participants who have a significant number of prior arrests (Babinski, Hartsough, & Lambert, 2001). For the FFT-only analysis, recidivism was operationalized as any sustained petition, as recommended by the Council of Juvenile Correctional Administrators (Harris, Lockwood, & Mengers, 2009).

The inclusion of parent and youth functioning in the outcome analysis was based on intra- and inter-personal factors known to affect participation in family treatment (Heatherington, Friedlander, & Greenberg, 2005). The conceptual definition of these mediating factors is problem behaviors (such as drug and alcohol use, violence, and self-harm), mental health symptoms (such as anxiety, psychosis, or suicidal ideation), and social problems (such as isolation or conflict) that could impede the expected progress of the family during treatment. These factors were measured to examine the magnitude of change in parent and youth behavior, symptom burden, and social functioning as a result of participation in FFT. They were operationalized as subscale score and total score on two measures: the Outcome Questionnaire and the Youth Outcome Questionnaire (Sexton & Alexander, 2004).

We planned to do an additional analysis to compare the cost-effectiveness of MST versus FFT in reducing recidivism among justice system-involved youth but were not able to collect all data prior to the deadline for this report. We anticipate providing the analysis to Contra Costa County at a later date. A reduction in recidivism can be conceptually defined as movement based on developmental progress among adolescent-limited or life-course persistent youthful offenders (Moffitt, 1993) toward desistance from crime (Laub & Sampson, 2001).

Epidemiological studies have been advanced due to the global burden of disease research (Institute for Health Metrics and Evaluation, 2018) and the consistent use of disability-adjusted life years (DALYs) in their analyses of morbidity outcomes. Similarly, researchers recommend the consistent use of quality-adjusted life years (QALYs) in cost-effectiveness studies of healthcare interventions (Drummond, Sculfer, Claxton, Stoddard, & Torrance, 2015; Eichler, Kong, Gerth, Mavros, & Jonsson, 2004). The QALY is the inverse of the DALY; the first accounting for years lived with a disability and the second accounting for years lived free of disease burden and in excellent health. Both are weighted for age and disease state and have been criticized as either advantaging or disadvantaging the young and those burdened with more than one disease state (Robberstad, 2009). We operationalized an equivalent measure for criminal justice research, an unweighted measure of desistance from crime or what Schawo et al. (2012) have termed criminal activity-free years (CAFYs). The CAFY can be used in a criminal justice cost-effectiveness or cost-benefit analysis, similar to the use of the QALY in healthcare research.

DATA COLLECTION AND ANALYSIS

Data Sources. Sources of data for this mixed methods evaluation varied, dependent upon whether they were used for process or outcome evaluation. For the MST versus FFT impact study, data were collected primarily from client files for all youth who were enrolled in either treatment and had a CALOCUS/LOCUS score available. For the FFT-only analysis, data were collected from the FFT clinicians and clinical supervisor about TOSO participants and the progress and outcomes, in terms of FFT services, for every referred youth. Covariate and outcome data about participants' prior- or post-FFT sustained petitions and detentions were collected from probation records and/or the County's court management computerized system. Data about fidelity to the FFT model were collected from FFT consultants with Functional Family Therapy, LLC and the California Institute

for Behavioral Health Solutions. Cost data were collected from the County's health services department and COFY.

Data Collection Methods. Data were collected by probation staff, COFY clinicians, and the MIOCR grant manager. The primary method of data collection was facilitated by the use of an Excel workbook with 20 individual worksheets. The workbook was shared via a secured connection. The tab and headings for each worksheet were color-coded to make it easier for probation or COFY staff to locate the datasheet which they were responsible for completing. Each month, probation staff entered information about the history of involvement with the court and probation for any new referrals; probation staff also located any sustained petitions or detentions for TOSO participants. Probation staff were also responsible for entering demographic information. Data collected and entered by COFY clinicians included open date, first and last session date, number of sessions, exit date, reason for exit, and status at closing. COFY staff were also responsible for entering pre-post school and employment information, and results of the FFT-specific assessments (OQ, YOQ, YOQ-SR, COM-A, COM-P, and FSRs). Additional datasets were maintained by the project evaluator, without any personal identifiers, for the MST versus FFT and the cost analyses.

Measurement Tools

CALOCUS or LOCUS. Level of care required to safely maintain the youth in the family and the community was measured using the Child and Adolescent Level of Care Utilization System (CALOCUS) for youth aged 12-18 and the Level of Care Utilization System (LOCUS) for youth older than age 18 (American Association of Community Psychiatrists, 2010). The CALOCUS/LOCUS was completed by the FFT or MST clinician after conducting an initial assessment to determine the appropriateness of family therapy delivered in the home. The purpose of the tool is to determine the level of service intensity needed to stabilize a client; it is also used by all Contra Costa County behavioral health clinicians in order to consistently allocate scarce mental health resources across the continuum of care (Wilbur, 2016).

A workgroup associated with the American Academy of Child and Adolescent Psychiatry, along with members of the American Association of Community Psychiatrists, who developed the LOCUS, developed the CALOCUS, also known as the Child and Adolescent Service Intensity Inventory (CASII) (Pourat, Zima, Marti, & Lee, 2017). The CALOCUS (American Association of

Community Psychiatrists, 2010) assists the clinician to identify protective and risk factors in the following 6 domains: 1) risk for harm, 2) functional status, 3) co-morbidity, 4) recovery environment (with two subscales: environmental stress and environmental support), 5) resilience and treatment history, and 6) treatment acceptance and engagement by the child and by the parent/guardian (see appendix B). The clinician assigns a score for each domain, from 1-5, with a higher score indicating need for a higher level of care. The clinician uses the higher of the two scores for the child/parent treatment acceptance/engagement domain and then sums them to obtain a composite score (American Association of Community Psychiatrists, 2010).

The LOCUS tool is used in much the same way as the CALOCUS (Wilbur, 2016). The domains include: 1) risk for harm, 2) functional status, 3) medical, addictive, and psychiatric co-morbidity, 4) recovery environment (with two subscales: environmental stress and environmental support), 5) treatment and recovery history, and 6) engagement. The clinician sums scores for each domain to arrive at a composite score. If risk of harm is scored as 4 (serious potential risk for harm), it is recommended that the child or adult receive services categorized as level of care 5 (see appendix C); if risk of harm is scored as 5 (extreme potential risk for harm), the clients should receive services categorized as level of care 6 (Wilbur, 2016). Clinicians determined the appropriate level of care as part of an assessment to determine eligibility for FFT or MST.

The reliability and validity of the CALOCUS/CASII was tested in four states, including the state of Hawaii who uses the tool throughout its entire Medicaid mental health care system (Fallon et al., 2006). Intraclass correlation coefficient (ICC) ratings for reliability among psychiatrists using the tool ranged from 0.73 to 0.93 for the subscale scores and 0.89 for the composite score. Among masters and doctoral level clinicians, the range of ICCs for subscale scores was 0.57 to 0.95, while the ICC for the composite score was 0.93. ICCs were lowest for the subscale score related to child acceptance of treatment (Fallon et al., 2006). ICC scores ranged from 0.16-0.72 in a reliability study of the LOCUS (Sowers, George & Thompson, 1999).

Validity of the CALOCUS/CASII was examined by comparing it to two other validated measures of a child's level of care need, the Child and Adolescent Functional Assessment Scale (CAFAS) with or without the Child Global Assessment Scale (CGAS) (Fallon, 2006). Pearson correlation coefficients (Pearson's r) were calculated to identify how similarly the three tools performed in identifying risk

scores for youth. There were low to moderate Pearson's r scores for most CALOCUS/CASII subscale scores and a 0.62 (low) correlation between the composite scores for the CALOCUS/CASII and the CGAS. Researchers sampling from the Tennessee system of care found similarly low subscale and composite score Pearson correlation coefficients (0.21 to 0.61) when comparing the CALOCUS/CASII and the CGAS (Pumariega, Millsaps, Moser, & Wade, 2014). No validity study of the LOCUS could be found for this report.

J AIS. The Juvenile Assessment and Intervention System (JAIS) was used to evaluate the risk level of participants (National Council on Crime and Delinquency [NCCD], n.d.). JAIS category was determined by the youth's probation officer at the time of court-ordered probation; JAIS was re-evaluated if youth committed another offense or when a placement change was indicated. JAIS categories reflect the assessment of youth/family needs and strengths as well as risk for ongoing antisocial behaviors (NCCD, n.d.). JAIS was obtained for each youth based on the closest date of assessment or re-evaluation prior to enrollment in FFT or MST.

The JAIS was developed from the Correctional Assessment and Intervention System (CAIS) for adults (NCCD, n.d.). The JAIS categories for risk level are high, moderate, or low. The risk assessment is based on answers to questions about youth characteristics in the following domains: history of delinquency, mental health status, health status, family and peer relationships, youth attitudes and values, relations with peers and family, general social skills and vocational abilities, performance in school, and the presence of substance misuse or abuse. Additional questions are included in the JAIS tool for girls, in order to identify "risk to public safety" (Glesmann & Irvine, 2014, p. 12). According to authors associated with the NCCD, the risk assessment embedded in the JAIS is based on the Juvenile Sanctions Center (JSC) risk assessment instrument (Baird, Healy, Johnson, Bogie, Dankert, & Scharenbroch, 2013).

Reliability and validity data about the JAIS could not be located for this report. However, Baird et al. (2013) evaluated the psychometric properties of the JSC, comparing it to 8 other risk assessment instruments. The JSC is an 8-item tool for girls and a 10-item tool for boys. Risk is determined based on information about the youth's age when first referred to juvenile court, history of disciplinary actions and attendance at school, parent's and/or sibling's criminal behavior, relationships with peers, and use or abuse of substances (Baird et al., 2013).

Reliability, measured as percent agreement between juvenile justice workers and between workers and researchers using the JSC was good at 92% for boys and 83% for girls (Baird et al., 2013). Intraclass correlation coefficients (ICC) were also calculated to determine how reliably the JSC instrument was being used across a large sample of youth; correlation between risk level of youth was 0.90 for risk category and 0.92 for risk score for boys and 0.74 and 0.89, respectively, for girls (Baird et al., 2013). Confidence intervals were not reported for these ICCs so it is not possible to determine whether reliability can be described as excellent or good, based on this measure (Koo & Li, 2016). Cohen's kappa (κ) statistic, another measure of interrater reliability was 0.78 for boys and 0.68 for girls (Baird et al., 2013). A kappa statistic of 1 is perfect agreement and a κ in the range of 0.61 – 0.80 indicates substantial agreement (McHugh, 2012).

Baird et al. (2013) also examined the validity of the JSC risk instrument. They identified the ability of the tool to correctly predict recidivism for youth with different measured levels of risk. After risk determination was made, 19% of low-risk boys and 24% of low-risk girls were adjudicated for a new offense within 12 months of risk prediction. Forty-eight percent of moderate-risk boys and 29% of moderate risk girls and 64% of high-risk boys and 58% of high-risk girls were adjudicated for a new offense within 12 months as well. Lastly, area under the curve (AUC) scores were determined to examine how well the instrument was able to distinguish between true positives (i.e. youth with persistent antisocial behavior who scored as high risk) from false positives (i.e. youth who desisted from antisocial behavior but scored as high risk). An AUC score of 1 means the predictive ability of the test is perfect and an AUC of 0.05 means the prediction was no better than it would be by chance; clinically significant predictive tests typically obtain an AUC between 0.70-0.80 (Youngstrom, 2014). The AUC for the JSC for boys and girls was 0.68 (Baird et al., 2013). The developers recommend that users of the JSC and JAIS risk instruments conduct regular validation studies to ensure that the prediction of risk is appropriate, given the geographic hazards in the local population (Howell & Lipsey, 2004; National Council on Crime and Delinquency, n.d.).

INDIVIDUAL, FAMILY, AND COMMUNITY RISK. Factors known to predict juvenile recidivism were also measured for the evaluation of FFT youth and family outcomes. We collected information about youth's prior involvement with probation, including age of first sustained petition, number of prior detentions and age of first detention, and number of probation violations filed with the court (Cottle, Lee & Heilbrun, 2001; Mallett, Fukushima, Stoddard-Dare, & Quinn,

2013). Data about the youth's individual, family, and social determinants of risk for recidivism at time of referral to FFT included age, gender, ethnic/racial group, geographic area of the county, employment, and school attendance, measured as the number of days attended during the prior 4 weeks (Trulson, Marquart, Mullings, & Caeti, 2005; Unruh, Gau, & Waintrup, 2009; Wolff, Baglivio, & Piquero, 2017). We also obtained information about the source of referral, including YOTP, GIM, the ranch, CSEY-service providers, placement probation officers, and community provider. Last, for the MST versus FFT sample, we obtained information about the youth's mental health diagnosis, entered into their clinical record at start of treatment.

A number of FFT-specific measures were used to assess youth and family engagement, clinician treatment adherence and fidelity to the FFT model, and therapeutic alliance between youth/family and clinician. These included the following measures.

ENGAGEMENT IN TREATMENT. Engagement of youth and family in treatment is measured based on the guidelines established by the developers of the FFT model (Sexton & Alexander, 2004). Accordingly, engagement is reflected in the number of referral cases who were contacted within 2 days of referral; the number of cases who were seen for the first session within 7 days of first contact; the number of cases who were seen for a second session within 7 days of the first session; and, the number of cases who were seen for a third session within 7 days of the second session. Sexton and Alexander (2004) suggest that FFT clinicians would be expected to engage at least 78% of families referred for treatment.

TREATMENT COMPLETION. Treatment completion is defined by Sexton and Alexander (2004) as the total number of days the youth/family received treatment, calculated as date the case was opened to date it was closed. Completion is defined as a total of between 60-180 days from open to close date (personal communication, March 13, 2015, Pam Hawkins, Senior Associate, California Institute for Behavioral Health Solutions).

TREATMENT ALLIANCE. Alliance is measured with similar tools, a Family Self Report (FSR) and Therapist Self Report (TSR). They are completed by everyone present after the first and second session during every phase of treatment for a total of at least 6 FSRs and 6 TSRs. The FSR is a 7-item tool, using a 7-point Likert scale, including questions such as 'Overall, how confident or

hopeful are you that your family will get better?’ and ‘Overall, how much do you approve or disapprove of the way your therapist is treating your family?’ The TSR is 6-items. Both are summed (for a range from 7-49) and a higher scores indicates greater alliance between family and therapist. Developers of the FFT intervention expect that therapists obtain all FSRs and compare them for agreement but set no benchmarks for total or average scores (personal communication, March 13, 2015, Pam Hawkins, Senior Associate, California Institute for Behavioral Health Solutions).

FIDELITY TO THE TREATMENT MODEL. Fidelity to the model for delivery of FFT was measured using qualitative and quantitative information. These data included meeting schedules and minutes relative to FFT training and consultation, supervision of clinicians, staffing of cases, review of clinical and case management records, and clinical coordination with referral source and other agencies serving the family and/or youth as outlined in the FFT training manual (Sexton & Alexander, 2004). Fidelity was also measured by collecting data about the results of audits of clinical files for adherence to MediCal-required assessments, treatment plans, and notes.

Tri-Yearly Program Evaluation (TYPE) reports, produced by analysts with Functional Family Therapy, LLC and the California Institute for Behavioral Health Solutions were reviewed to determine whether treatment pacing, the number of open cases, the number of cases successfully terminated, the number of failures and the active caseload for each trimester was adequate, according to FFT guidelines (California Institute for Behavioral Health Solutions, n.d.). TYPE report data included information about each individual clinician’s performance (attendance at consultation meetings, caseload, and supervisor rating of clinical competence as well as fidelity to the FFT model) and the overall team performance. In addition to the measurement of engagement in treatment and treatment completion, TYPE reports provided information about four other measures of adherence: 1) utilization, 2) supervisor rating of each clinician’s clinical competence, 3) supervisor rating of each clinician’s adherence to the FFT model, and 4) completion of pre-post treatment measures.

FFT UTILIZATION. Utilization is defined as the individual and team caseload for the select time period, given the expected caseload for each clinician and for the team. A part-time clinician is expected to have an active caseload of 5 families at any point in time. Utilization is calculated as the number of active cases in the case management system on the last day of the reporting period as a

percentage of the expected caseload. It is expected that individuals/teams will have a utilization rating between 75-100%.

CLINICAL COMPETENCE AND FFT ADHERENCE. Competence in delivering FFT is measured by the clinical supervisor, using a FFT-specific rating system. Competence is assessed based upon the clinician's weekly case presentation and is measured from 0-5. Average score equal to or greater than 3 is expected. Adherence to the FFT model is also measured by the clinical supervisor, using a rating system measured from 0-5. Adherence refers to those activities that occur outside of the actual session, such as timely response to referral, treatment pacing, maintenance of clinical records, and completion of required clinical assessments. Average score equal to or greater than 4 is expected.

OUTCOME QUESTIONNAIRE (OQ). The OQ is completed by any parent or guardian in the household during the first and last session of FFT. The questionnaire consists of 45 Likert-scale items measuring three areas of functioning: symptom distress (SD), interpersonal relations (IR), and social role performance (SR). Nine items, measuring satisfaction with quality of life, are reverse scored; there are also three critical items that ask about suicide, drug or alcohol abuse, and potential for violence. Cut scores have been developed to identify results that are consistent with those obtained by clinic patients versus community samples. Subscale cut scores are 36 for SD, 15 for IR, and 12 for SR; the cut score is 64 for total OQ score. Pre-post improvement is indicated if the individual with a clinical cut score has a post-treatment score in the range of community samples and/or if there is a 15 point reduction in total score (Wells, Burlingame, Lambert, & Hoag, 1996). Clinicians are asked to examine the OQ responses during the treatment session and further investigate if the parent has answered positively to one of the three critical items (California Institute for Behavioral Health Solutions, n.d).

Reliability of the OQ was assessed using the test-retest method, examining correlations between repeated measures using Pearson's product movement correlation coefficient (Pearson's r). Perfect positive correlation results in a Pearson's r equal to 1, a perfect negative correlation results in a Pearson's r equal to -1 and no correlation whatsoever is equal to 0 (Rodgers, & Nicewander, 1988). A Pearson's r of 0.10 is considered small, 0.30 is medium, and 0.50 is large (Cohen, 1988). Pearson's r for subscales were calculated as SD=0.78, IR=0.80, and SR=0.82 and Pearson's r for total score

was 0.84 (Burlingame, Lambert, Reisinger, Neff, & Mosier, 1995).

Reliability was also assessed using Cronbach's alpha to examine how consistently the subscale items measured SD, IR, and SR and how well the sum of scores reflected a true score for global functioning. Cronbach's alpha scores of 0.70 and above are desirable (Nunnally, 1978). Cronbach's alphas were 0.92, 0.74, and 0.70 for SD, IR, and SR, respectively, and 0.93 for total score. Validity of the OQ was assessed by comparing results of the OQ in a normative sample of over 1,000 adults to results obtained using several measures of anxiety, depression, interpersonal relations, and social functioning. Cronbach's alphas for total scores ranged from 0.44-0.88. (Burlingame et al., 1995).

YOUTH OUTCOME QUESTIONNAIRE (YOQ). The YOQ is completed by the parent on behalf of the youth during the first and last treatment sessions. Youth also complete a YOQ-self-report (SR). The YOQ was developed as equivalent to the OQ, also used to measure change during treatment. Both the YOQ and the YOQ-SR are parallel, 64-item questionnaires with 5-point Likert scale responses asking about the child or adolescent's behavior during the past week. Response options include 'never or almost never,' 'rarely,' 'sometimes,' 'frequently,' and 'almost always or always.' Responses are totaled, with items that assess protective factors reverse scored. Subscales are used to assess intrapersonal distress (ID), somatic symptoms (S), interpersonal relations (IR), social problems (SP), behavioral dysfunction (BD), and critical items (CI; assessing for problems such as suicidal ideation or psychosis). Cut scores for subscales are 16.4 for ID, 5 for S, 4.4 for IR, 3 for SP, 12 for BD, and 5 for CI with a total score cut point of 46 for the YOQ and 47 for YOQ-SR (California Institute for Behavioral Health Solutions, n.d).

The total score ranges for the YOQ and YOQ-SR are from -16 to 240; the total score reflects the degree of distress among youth at start and end of treatment. A change of 13 points and scores below clinical cut points are used a reliable measure of improvement over time in treatment (Dunn, Burlingame, Walbridge, Smith & Crum, 2005). Reports are obtained from both parents and youth to address the possibility that the youth has not disclosed problem behaviors or critical symptoms to the parent. Older youth, girls, and youth with antisocial behaviors are likely to self-report more problem behaviors than their parents report on their behalf (van der Ende & Verhulst, 2005). Dunn et al. (2005) examined YOQ and YOQ-SR results for a sample of 296 youth in the community, comparing them to 5,132 YOQ and 4,131 YOQ-SR results from clinic patients and

parents. Reliability for the YOQ and YOQ-SR was similar to that of the OQ, with Cronbach's alphas of 0.92 for the total score for both measures. Pearson's r for test-retest reliability were 0.84, 0.80, and 0.91 for the total sample, parent sample-only, and youth sample-only, respectively (Dunn et al., 2005).

Validity of the YOQ and YOQ-SR was tested using receiver operating characteristic (ROC) analysis; a ROC analysis produces sensitivity and specificity proportions which are graphed on the x- and y-axis (Hajian-Tilaki, 2013). Examining the area under the curve (AUC), Dunn et al. (2005) determined whether a randomly selected YOQ score belonged to a youth from the community or the clinic sample. An AUC of 0.50 indicates the result of the analysis were no better than would be obtained by chance while an AUC of 1.0 indicates 100% predictive ability (Youngstrom, 2014). The AUC for the YOQ was 0.84 (95% confidence interval of 0.82-0.87) and for the YOQ-SR was 0.79 (95% confidence interval of 0.74–0.84). These results indicate that the YOQ was able to discriminate between community and clinic samples 84% of the time and the YOQ-SR was able to discriminate 79% of the time (Dunn et al., 2005).

McLendon et al. (2011) examined the YOQ and YOQ-SR, comparing the tools to two other commonly used measures of change during treatment. Using hierarchical linear modeling, they found that a child or adolescent's wellbeing measured using the YOQ was the best indicator of change among the three tools, consistent with clinicians' rating of both youth improvement and deterioration over time. During 25 weeks of observation, for youth who showed improvement, every one unit increase in improvement each week ($\text{intercept} + \log_e + \text{week} + 1 * \text{slope}$) was reflected by a corresponding 2.29 change in total YOQ score. Changes in the total score on the Child Behavior Checklist (CBCL; Achenbach, 1991) and the Behavior Assessment System for Children-2 (BASC-2; Reynolds, & Kamphaus, 2004) were 1.01 and 1.10, respectively (McLendon et al., 2011).

ADHERENCE TO COURT-ORDERED PROBATION. For purposes of this project, we measured adherence of youth and family to probation-required and court-ordered activities. We measured adherence as an absence of petitions to the court for technical violations of juvenile probation, submitted according to California's Welfare and Institutions Code, section 777 (State of California, 2018). Though the reliability and validity of adherence to juvenile probation requirements as a measure of protective factors has not been established in the literature, its use as

an additional measure of risk for recidivism has been discussed (Council of State Governments Justice Center, 2014).

CLIENT OUTCOME MEASURE (COM). At completion of FFT, all present at the last session are asked to complete a measure of self-reported improvement (California Institute for Behavioral Health Solutions, n.d.). The Client Outcome Measure-Adolescent (COM-A) and the Client Outcome Measure-Parent (COM-P) consist of 6-items using a 6 point Likert-scale (0='Things are worse', 1='No change', 2='Only a little better', 3='Somewhat better', 4='A lot better', and 5='Very much better'). The 6 items ask about a) overall level of family change; b) change in communication skills; c) change in adolescent behavior; d) change in parenting; e) change in parental supervision; and f) change in family conflict. The clinician completes a Therapist Outcome Measure (TOM) that includes the same 6 items as the COM-A and COM-P. No reliability and validity information could be located for the COM tools; the family is considered to have improved as a result of FFT completion if COM-A, COM-P, and COM-T scores are consistently 3, 4, and 5 (personal communication, July 30, 2018, Lori Ford, Associate, California Institute for Behavioral Health Solutions).

RECIDIVISM. We measured recidivism for the FFT and MST participant sample using reports of arrest at least 6 months after completion of family therapy as measured by any arrest with a recorded date in the Criminal Offender Record Information (CORI) section of the California Law Enforcement Telecommunications System (CLETS). Data were obtained and maintained by probation staff, and collected only for youth with unsealed juvenile records, consistent with guidelines for researchers (California Department of Justice, n.d.). Arrest records obtained through CLETS are an adequate measure of recidivism, especially for locating arrests throughout California and in other participating states (Allen & Grassel, 2017). Recidivism for the FFT-only sample was measured using sustained petitions filed at least 6 months after completion of FFT, consistent with the definition of recidivism recommended by the California State Attorney General (2014).

TREATMENT COSTS. Last, we collected data about the cost of providing family treatment to probation referred youth in Contra Costa County. These were the cost of services billed to the county, training and consultation costs, and costs associated with County personnel working on behalf of the project but not grant-funded. The purpose of collecting data about treatment costs was

to provide an estimate of the cost of TOSO were it not funded by the MIOCR grant.

Data Limitations

As this evaluation was designed to examine the effectiveness of family therapy in reducing recidivism for youth transitioning back to the community after in-custody treatment, limited data were available for family composition and neighborhood characteristics. These data are critical to any determination of predictors of rearrest, given participation in a social-ecological model of care (Bronfenbrenner, 1979). Assessment of risk and protective factors as well as anticipated responsiveness to treatment (Andrews, Bonta, & Wormwith, 2006), was also limited by the data available. Our analysis of baseline youth characteristics associated with rearrest was further hampered by the use of risk and responsiveness measures (the JAIS and CALOCUS/LOCUS) that have not demonstrated reliability and validity in predicting desistance from or ongoing participation in crime among youth involved in the justice system.

The data used for the evaluation of impact for FFT alone and for MST versus FFT were limited due to the lack of longitudinal data. Though we had reliable and valid pre-post measures of change, measurement of recidivism was restricted due to the limited follow up period for the FFT sample and the different follow up period for the MST sample. The varied observation periods could be controlled for with the use of advanced statistical tests; however, data for the outcome variables (rearrest and sustained petitions post-treatment) for all samples were not normally distributed, limiting our ability to make inferences about recidivism. We also collected information about only the first arrest post-MST or FFT further limiting a longitudinal view of recidivism or desistance from crime.

Additionally, the data used to measure recidivism among MST and FFT youth were obtained using the Criminal Offender Record Information (CORI) section of the California Law Enforcement Telecommunications System (CLETS); these data provide information about arrests outside of Contra Costa County as well as any arrest within the county. We used arrest data to capture events within the 18 month period of follow-up rather than missing observations while waiting for the outcome of the arrest. However, arrest data are limited because they may or may not depict an actual return to crime but rather to increased police presence in high crime neighborhoods, consistent with

the broken windows theory of crime reduction (Corman & Mocan, 2005). For the FFT-only analysis, we used sustained petitions post-treatment as the measure of recidivism. These data more accurately measure return to criminal behavior since they are equivalent to a finding of guilt in the adult criminal justice system. However, the use of juvenile court records is a limited measure of further criminal involvement since youth may be arrested but not charged, arrested but diverted to receive services, arrested and formally charged, adjudicated and make a plea agreement, or adjudicated and returned to custody (Cottle et al., 2001). Ideally, criminal activity should have been measured using multiple sources, including self-report, since a substantial proportion of youth participate in criminal behavior but are never arrested (Farrington et al., 2003).

Another limitation of the data used to evaluate the impact of FFT on recidivism was related to policy changes during TOSO implementation. One target population for TOSO included commercially sexually exploited youth (CSEY). The inclusion of CSEY in the targets for treatment was based on a GIM program evaluation (Shade & Ceric, 2017) in which we examined recidivism for the youth remanded to the program since its inception. We found GIM participants identified as CSEY were at greater risk for a rearrest and a return to custody. However, as of 2016, youth should no longer be arrested for prostitution-related offenses; rather, in compliance with federal law (PL 113-183, the Preventing Sex Trafficking and Strengthening Families Act) and state law (Senate Bill 794), child welfare social workers and juvenile probation officers are required to identify youth who are or are at risk of becoming CSEY so they can be diverted from custody. SB 794 requires that youth with 2 or more risk factors for CSEY be screened and referred to the county's child welfare department (State of California, 2016). As a result of these policy changes, participants were less likely to be identified as CSEY or to be referred specifically for involvement in sex work. We hope that the youth we anticipated targeting when we submitted the MIOCR grant were diverted from the criminal justice system and in receipt of supportive services in order to reduce their risk for CSEY, but we had no data to evaluate whether this occurred.

Data Analysis for Impact Evaluation

For the MST versus FFT analysis, we used CALOCUS/LOCUS categories (Fallon et al., 2006) to create a comparative sample with similar baseline risk. Since MST had been delivered to justice-involved youth from 2013 to the present, and FFT was implemented in October of 2015, we had to

control for time as an influence on rearrest for youth among the two groups. We did this by collecting rearrest data for youth 6 months after they had completed MST or FFT, measuring time to rearrest between 6-18 months after family therapy. We also removed youth from the analysis if they were observed post-treatment for longer than 3.5 years. Last, we used a time-to-rearrest design in order to compare the hazard for rearrest among MST versus FFT participants. This statistical procedure is appropriate when it is possible to average time-to-event HRs during a period of follow up (Hernán, 2010).

PROCESS EVALUATION RESULTS

Coordination of Services. As a result of process evaluation during monthly case review meetings, it became apparent that the team needed to improve the coordination of services prior to making a TOSO referral. Several times, family members expressed confusion due to inconsistent messages from probation officers and mental health liaisons. There were some cases in which the family was court-ordered to complete family therapy and had been inadvertently referred to Multidimensional Family Therapy (MDFT), provided by a community based organization in Contra Costa County, as well as MST and/or FFT.

The COFY staff were able to address multiple referrals to MST and FFT but had difficulty engaging families if they had been referred to MDFT as well as one the therapies delivered by COFY. To address this problem, the TOSO team delivered numerous messages, in writing and in person, to inform probation officers and in-custody treatment staff about the TOSO referral process. The liaisons also appointed a lead among the three who took responsibility for maintaining a central list of youth exiting the in-custody treatment programs and facilitating regular meetings between probation referral sources and FFT clinicians. Additionally, the FFT supervisor and clinicians worked one day a week out of the three field probation offices to improve communication and coordination of services.

Timeliness of Treatment. The TOSO team worked hard to respond quickly to referrals and engage with youth and families while the youth was still in custody. This was difficult, requiring excellent communication between in-custody staff, mental health liaisons, and FFT clinicians. The process seemed to work most smoothly when the in-custody mental health service provider connected

directly with the FFT clinician and supported the youth as they made the transition from in-custody to community treatment. The FFT teams' timely response to referrals and in-custody visit with youth and families was impressive given the services were grant-funded; no service delivered while the youth in custody can be billed to MediCal. Though youth eligible for MediCal are entitled to diagnosis and treatment of mental illness, as well as service coordination and intensive community-based mental health services through the Early and Periodic Screening, Diagnostic and Treatment (EPSDT) program, the county will not be reimbursed from the federal government for EPSDT services provided to youth while they are still in custody (California Department of Health Care Services, n.d.).

Responsiveness to referrals was additionally impressive, given the periods of instability and transition in the program. Working with the highest risk youth returning from out-of-home placements was difficult and there were several instances when a clinician experienced vicarious or secondary trauma in the line of service (Pearlman & MacIan, 1995). Some youth were at serious risk of violence once released from custody and therapists were challenged to identify ways the family might be able to keep the youth safe from harm. In one instance, the clinician assisted a parent to negotiate with the court to allow the youth to return to his South American country of origin after he was threatened with death by other youth, even though the client had been living in the US for years. The clinician successfully supported the family to protect this young person from harm though it meant the case was among those closed unsuccessfully in FFT. In two other incidents, youth successfully completed FFT but were then shot and killed, greatly impacting the clinicians who had worked with the family.

The FFT clinicians recognized the benefit of engagement of youth and families while the youth was still in custody during the last quarter of FFT (April-June, 2018). At that time, both MDFT and MST clinicians were not able to serve the number of youth and families on their waitlists. Anticipating the end of the MIOCR-funded activities, the FFT supervisor agreed to accept referrals from the two program waitlists. Some of these families had been waiting 3-4 months for treatment. At least 10 of the youth and families referred failed to engage in treatment, likely due to the lack of timeliness of treatment.

Fidelity to the FFT Model. An important component of the evaluation were regular clinical

conferences and reviews of clinical records to ensure FFT was being delivered as designed. The consultants from Functional Family Therapy, LLC (n.d.) who were responsible for initial training and oversight and the associates from the California Institute for Behavioral Health Solutions (n.d.), who were responsible for assisting with FFT adherence while sustaining the program, were central to this component of the process evaluation. The FFT clinical team members regularly entered data into the clinical services system, an online database used by consultants and the FFT supervisor to ensure individual clinician competence in FFT and team adherence to the model. Data were available to track number referred/number opened, open caseloads by individual clinician and team, dropout rates and reasons, treatment pacing, number of sessions and number in each phase of treatment, and completion status. Results of clinical assessments used for treatment planning and evaluation were also available for review. Clinicians consulted weekly during the first 18 months of implementation then supported the clinical supervisor in taking responsibility for evaluating adherence during the second year. Throughout implementation, California Institute for Behavioral Health Solutions associates provided additional consultation and produced the Tri-Yearly Program Evaluation (TYPE) reports, to support the FFT team in making improvements.

Throughout the TOSO project, we evaluated TYPE report data to examine trends in the development of FFT adherence to the delivery model. TYPE reports were produced three times a year, providing feedback to the team for fourth month periods of time; the first year, the initial report was for services delivered between October 12, 2015 and February 26, 2016 and the last report was for services delivered between January 1, 2018 and May 1, 2018. As expected, there were improvements seen in clinical competence and adherence to the FFT model over time with the exception of two quarters when they were rebuilding the clinical team (see Table 2).

TABLE 2

Clinical Competence and FFT Adherence according to TYPE Reports

	FFT Utilization Rates	Average Treatment Pacing	Average Clinical Competence Scores	Average FFT Adherence Scores	Average Completed Measures
Oct 2015-Feb 2016	43%	12.26	3.24	4.95	65.5%
Jan 2016-May 2016	57%	16.22	3.91	5.52	66.25%
May 2016-Sept 2016	100%	11.57	3.65	4.42	92%
Sept 2016-Jan 2017	77%	11.62	3.52	4.81	94.5%
Jan 2017-May 2017	100%	13.09	4.38	4.62	85.25%
May 2017-Sept 2017	49%	11.45	2.85	4.63	82.3%
Sept 2017-Jan 2018	93%	15.26	2.95	4.37	95.75%
Jan 2018-May 2018	70%	12.55	3.7	4.64	95.5%
TOTALS	74%	13 days	3.53	4.75	85.5%

Note: Average treatment pacing calculated as average of three averages: days from referral to first session, days from first session to second session, and days from second session to third session

Clinical competence and adherence to the FFT model were also assessed using a measure of treatment alliance. Families are expected to progress through the phases of FFT when each family member is engaged, motivated, and aligned with the therapist (Sexton & Alexander, 2004). We examined the results of the Family Self-Reports (FSR) to determine treatment alliance. We did not have access to Therapist Self-Report (TSR) data; since we had only the 7-item FSRs, we averaged items scores rather summing the items and then comparing averages. Though data were not available during every phase of treatment, there was at least one FSR for every family opened in FFT.

Across the entire sample, the average score on youth FSRs was 6 ($M=5.86$, $SD=0.69$, $n=195$), the average score for mother (mother-figure) FSRs was 6 ($M=5.96$, $SD=0.56$, $n=173$) and the average score for father (father-figure) FSRs was 6 ($M=5.89$, $SD=0.64$, $n=66$). A Pearson correlation analysis, comparing the average FSR score for youth and for mothers, resulted in a correlation coefficient of 0.46, indicating a moderate effect size (Cohen, 1988). A significant positive correlation was observed between average of youth and average of father FSRs ($r_p=0.68$, $p < 0.001$) indicating a

large effect size. Mother and father FSRs were also significantly positively correlated ($r_p=0.55, p<0.001$) with a large effect size. These results suggest there was agreement across youth and parent(s) about the alliance with FFT clinician and that as FSR increased for one family member, there was an associated increase in FSR for the other family member(s).

An additional measure of clinical competence were the Client Outcome Measures collected at the last session and used to assess adolescent and parent self-report of improvement as a result of participation in FFT. The mean values for the Client Outcome Measure-Parent Report were all above 3 ('a little better'), the value established as a measure of client and family improvement (personal communication, July 30, 2018, Lori Ford, Associate, California Institute for Behavioral Health Solutions). The lowest mean value was for behavior (3.99) and the highest was for overall change (4.17). There was very little variability in the data (see Table 3).

TABLE 3
Mean Scores for the Client Outcome Measure-Parent Report (COM-P)

	<i>M</i>	<i>SD</i>	<i>n</i>	<i>SE_M</i>
Change	4.17	0.86	114	0.08
Communication	4.03	0.94	112	0.09
Behavior	3.99	1.00	113	0.09
Parenting	4.03	0.90	114	0.08
Supervision	4.04	0.95	114	0.09
Conflict	4.08	1.07	113	0.10

Mean scores on the Client Outcome Measure-Adolescent Report were all above 3 as well. Youth rated improvement highest in the area of conflict and lowest for communication (see Table 4).

TABLE 4

Mean Scores for the Client Outcome Measure-Adolescent (COM-A)

	<i>M</i>	<i>SD</i>	<i>n</i>	<i>SE_M</i>
Change	3.95	1.04	117	0.10
Communication	3.90	0.96	117	0.09
Behavior	3.91	1.18	116	0.11
Parenting	3.98	1.11	117	0.10
Supervision	3.98	1.11	117	0.10
Conflict	4.09	1.09	117	0.10

IMPACT EVALUATION RESULTS

MST and FFT Probation Youth

Initial MST and FFT Sample. The sample we obtained for comparison of MST versus FFT were 268 youth who were on probation; 174 (65%) were referred for MST during the years 2013-2017 and 94 (35%) were referred for FFT during the years 2015-2017. Both programs served youth not on probation (for FFT, 4% of their caseload was non-probation youth; this information was unknown for MST). The youth selected for the analysis were those for whom we had a CALOCUS/LOCUS determination of need and, in most cases, a JAIS determination of risk. This was done to ensure the MST and FFT samples were comparable. The sample consisted of primarily male (77%) youth between the ages of 12 and 19 with an average of 16.38 years of age ($SD=1.46$, $SE_M=0.09$, Range=12.39-19.77 years). Among the entire sample, 68% completed either MST or FFT. Seventeen youth (6%) completed MST and were subsequently referred to FFT. The average days in treatment were the same for both groups ($M=124.09$ for MST; $M=123.59$ for FFT) but there was greater variability in length of treatment among the FFT youth.

The MST and FFT participants were similar in terms of demographics, given the variables we were able to examine, except for age. As expected, the average FFT youth was older ($M=17.23$, $SD=1.26$, $SE_M=0.13$, Range=13.73-19.77); the average MST youth was 1.3 years younger ($M=15.92$, $SD=1.36$, $SE_M=0.10$, Range=12.39-19.61). Males made up 75% of the MST sample and 81% of the FFT sample.

We had information about the youth's mental health status, according to criteria referenced in the diagnostic and statistical manual of mental disorders (American Psychiatric Association, 2013) for 82% of the sample; there were differences between DSM disorders when comparing MST and FFT youth. More MST than FFT youth had no DSM diagnosis; 18 (13%) among MST and 3 (4%) among FFT. There were a greater number of youth with externalizing disorders such as oppositional defiant and attention deficit disorders among MST referrals (95, 68%) than FFT referrals (32, 41%). More FFT than MST youth had an internalizing disorder such as a depressive disorder, 29 (37%) for FFT and 22 (16%) for MST. A greater number of youth referred to FFT had a primary diagnosis of Posttraumatic Stress Disorder (PTSD); 15 (19%) among FFT and 5 (4%) among MST.

For a baseline risk comparison, we examined JAIS categories for those available (21% were missing for MST youth; 3% for FFT youth); we used the CALOCUS/LOCUS category as an additional measure of need. The majority of both samples were of moderate risk according to the JAIS (32% for MST and 47% for FFT). The majority determination of need differed between the two groups with 42% of MST youth requiring low-intensity community based services (14-16 CALOCUS total score) and 46% of FFT youth requiring medically monitored community based services (20-22 CALOCUS total score). This finding confirmed that TOSO was delivering services to youth of higher risk, greater needs.

An independent samples t-test demonstrated a statistically significant difference in total CALOCUS/LOCUS score for MST versus FFT youth (MST: $M=19.13$, $SD=3.94$; FFT: $M=22.05$, $SD=3.78$; $t=-5.87$, $p<.001$, $d=0.760$; degrees of freedom=266). Cohen's d , an indication of the magnitude of the difference between the two groups, indicated a fairly large difference in needed level of service for MST versus FFT youth (Cohen, 1988). Given the results of the t-test, we removed 63 low-needs youth from the MST group and re-examined this final sample to identify any differences between MST and FFT participants.

Final MST and FFT Sample

The final sample for the comparison of rearrest for MST versus FFT participants on probation consisted of 203 youth, 110 (54%) were referred for MST during the years 2013-2017 and 93 (46%) were referred for FFT during 2015-2017. One hundred thirty one (64%) of the youth completed

MST or FFT treatment, with 7% receiving at least one session of both interventions.

There was little difference between the two groups in terms of gender, but, as expected, FFT youth were older than MST youth (see Table 5 and Table 6). MST youth, on average, were 1.4 years younger than FFT youth. Again, the average length of time in treatment was the same for both groups but there was greater variability in length of treatment for the FFT group ($M=122.40$, $SD=48.82$ for MST and $M=123.41$, $SD=60.71$ for FFT).

TABLE 5
Gender for MST versus FFT Youth

	MST	FFT
Males	84 (76%)	75 (81%)
Females	26 (24%)	18 (19%)

TABLE 6
Age for MST versus FFT Youth

	<i>M</i>	<i>SD</i>	<i>n</i>	<i>SE_M</i>
MST	15.86	1.33	110	0.13
FFT	17.26	1.24	93	0.13

We reexamined descriptive statistics for mental health disorder among the MST and FFT sample. Fourteen (7%) youth had no diagnosis and 19 (9%) had missing data. The majority of both MST and FFT youth were diagnosed with an externalizing disorder but there were a greater number with an internalizing disorder and posttraumatic stress disorder (PTSD) among the FFT sample. Youth with more than one disorder were equally represented in the two treatment groups (see Table 7). We further examined the difference in expected frequencies for each mental health diagnostic category across the two treatment groups. There were no statistically significant differences between MST versus FFT group for youth with more than one disorder but for all other category of disorder, there were statistically significant differences between observed and expected frequencies. There were statistically significantly fewer observations than expected in the MST group for internalizing disorders ($\chi^2(1)=8.10, p=0.004$) and for PTSD ($\chi^2(1)=8.00, p=0.005$) and fewer observations than expected in the FFT group for no disorder ($\chi^2(1)=4.57, p=0.033$) and externalizing disorders

($\chi^2(1)=9.91, p=0.002$).

TABLE 7

Mental Health Disorders for MST versus FFT

	MST	FFT
No diagnosis	11 (11%)	3 (3%)
Externalizing disorder (e.g. conduct disorder)	64 (66%)	33 (38%)
Internalizing disorder (e.g. depressive disorder)	11 (11%)	29 (33%)
PTSD	3 (3%)	15 (17%)
More than one diagnosis	8 (8%)	7 (8%)
Missing	13 (68%)	6 (32%)

The final MST and FFT youth also differed on risk and determination of level of need categories. The most common risk category for MST youth was low (30%) but the majority of FFT youth were determined to be of moderate risk (46%) (See Table 8). Twenty-eight youth (13.8%) had missing JAIS data. Risk level for the observations among MST youth was not statistically significantly different than what would be found by chance; however, among the FFT youth, results of a chi-square goodness of fit test were significant, $\chi^2(2)=9.27, p=.0010$. There were fewer observations than expected for high and low risk youth in the FFT sample.

TABLE 8

JAIS Categories for MST versus FFT Youth

	MST	FFT
Low Risk	33 (30%)	27 (29%)
Moderate Risk	29 (26%)	43 (46%)
High Risk	23 (21%)	20 (22%)

Note. Due to rounding errors, percentages may not equal 100%. JAIS category was missing for 25 (23%) of MST youth and 3 (3%) of FFT youth.

The majority level of care required for MST youth was category three, with 38% of the sample

determined to need high-intensity community based services. For FFT youth, the majority was category four, with 45% of the sample determined to need medically monitored community based services (see Table 9). These data depict the category of CALOCUS/LOCUS, not the total scores.

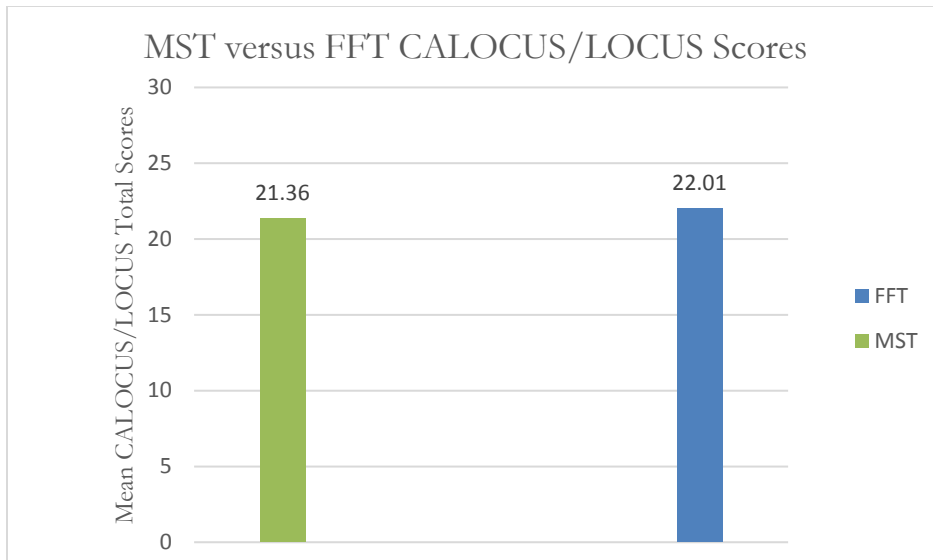
TABLE 9
CALOCUS/LOCUS Level of Need (Category) for MST versus FFT Youth

	MST	FFT
0: Preventive Services	0 (0%)	0 (0%)
1: Maintenance Services	0 (0%)	1 (100%)
2: Low Intensity Community Services	1 (17%)	5 (83%)
3: High Intensity Community Services	37 (64%)	21 (36%)
4: Medically Monitoring Community Services	42 (68%)	20 (32%)
5: Medically Monitored Residential Services	23 (35%)	42 (65%)
6: Medically Managed Residential Services	7 (64%)	4 (36%)

Note. Due to rounding errors, percentages may not equal 100%.

Baseline Comparison for MST versus FFT Youth. We further examined the difference in level of care needed for youth referred to MST and those referred for FFT, using a Welch's t-test to examine mean CALOCUS/LOCUS total scores for the two groups. Results of the t-test indicated there was no statistically significant difference in CALOCUS/LOCUS total scores (MST: $M=21.36$, $SD=3.05$; FFT: $M=22.01$, $SD=3.78$; $t=-1.33$, $p= .186$, $d=0.19$; degrees of freedom=175.97). The mean value for both groups was in the range of 20-22, CALOCUS/LOCUS determined level of need category 4; for medically monitored community based services (see Figure 1).

FIGURE 1. Mean CALOCUS/LOCUS Total for MST versus FFT Youth



Having determined that the two groups were comparable with regard to risk level and needs, we set out to determine the impact of the two interventions on recidivism, measured as rearrest 6-18 months after treatment with MST or FFT.

MST versus FFT Rearrests

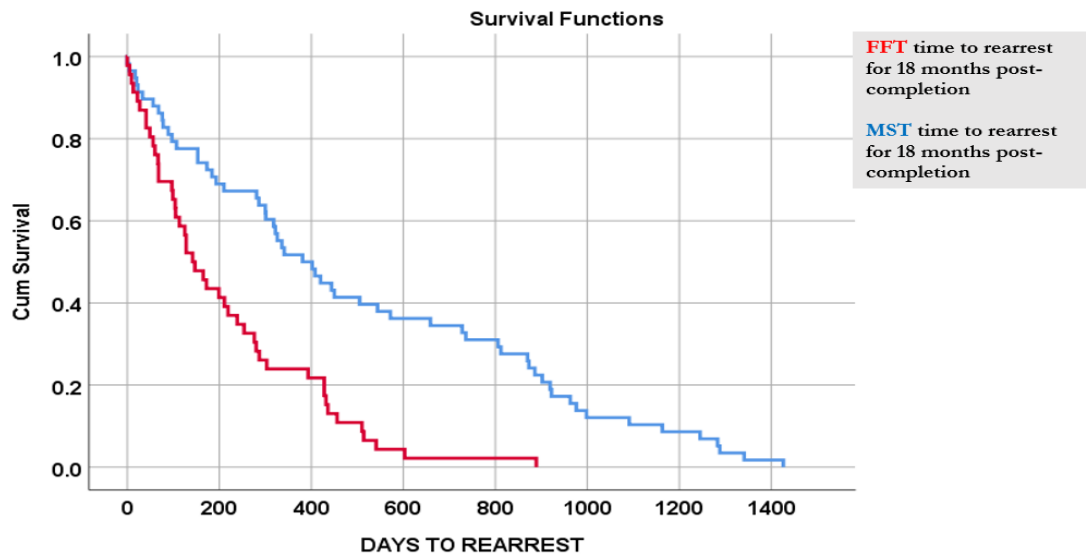
We examined the difference between rearrest 6-18 months after family therapy among the sample of MST and FFT youth probationers, removing 23 youth who were rearrested more than 18 months after treatment. We conducted a logistic regression to test the effect of family therapy (MST or FFT) as well as days in treatment, used as a ratio-level variable representing the dose of therapy, on the likelihood that youth will be rearrested. The benefit of a logistic regression, with a large enough sample and with the assumptions of the test met, is that multiple variables can be tested to determine the degree to which each predicts a binary outcome variable (Stoltzfus, 2011). Another benefit is that an odds ratio, a measure of effect size, can be determined.

We examined the rearrest data to ensure independence of observations and other assumptions for a logistic regression were satisfied. The linear relationship between the continuous predictor (days in treatment) and the logit odds was tested using the Box Tidwell test and was found to be non-significant, indicating that the assumption of linearity of independent variables and log odds outcomes was not violated (Stoltzfus, 2011). The overall logistic model was statistically significant,

$\chi^2(2)=9.786, p=0.008$. However, the model explained only 7.1% (Nagelkerke's R^2) of the variance in rearrest among the sample. It also only correctly classified 58.3% of cases. Results indicate that FFT youth were 1.3 times (95% confidence interval of 0.7-2.345) more likely to be rearrested than MST youth. A greater number of days in treatment was associated with a reduced likelihood of re-arrest.

To measure the impact of MST versus FFT on a reduction in recidivism following treatment, we also conducted a Kaplan Meier survival analysis, comparing days to rearrest for MST to FFT. A survival analysis is a statistical tool used to examine longitudinal data about a specific event; in this case the event was rearrest (Aalsma, White, Lau, Perkins, Monahan, & Grisso, 2015). A survival analysis is preferred when data are censored; that is, when observations between subjects varied by time and when, at any one point in time the probability of the event can be determined as having occurred (i.e. the youth was rearrested), has not yet occurred, or is unknown (i.e. the case was censored). An assumption of a survival analysis is that cases who remain event-free are still at risk of experiencing the event. Cases are right-censored if they are lost to follow up or have not experienced the event prior to the end of the observational period; their continued risk for the event is accounted for in the model. Cases are left-censored if they were diagnosed or began treatment before they were under observation. Censored cases and the pattern of censorship should be similar among the groups being compared (Aalsma et al., 2015). Censored cases were similar among all MST and FFT youth. The overall cumulative survival distributions are graphed in Figure 2. Results of the Kaplan Meier survival curves clearly demonstrated that there were less cumulative days-to-rearrest for youth who completed FFT than youth who completed MST.

FIGURE 2. Kaplan-Meier Survival Curves for Days to Rearrest for FFT and MST Youth



MST versus FFT Rearrests (2016-2018)

To further examine the risk for rearrest among MST versus FFT, we compared rearrests for only those youth who entered treatment for MST or FFT between January 2016 and October 2017, ending treatment between March 2016 and February 2018. This reduced the number of youth in FFT who were treated during an unstable period of start-up and balanced the numbers in the two groups. To address the concern that the two groups might not be comparable in risk, we used the JAIS to determine risk for recidivism and the CALOCUS/LOCUS score to determine level of needed service. As was the case in the entire sample, risk level was greater for youth referred to FFT. However, the average CALOCUS/LOCUS score for MST and FFT youth were the same.

We examined the demographic characteristics of the 57 MST and the 55 FFT youth who completed the programs in 2016-2018 and were rearrested. There were 14 MST youth and 27 FFT youth who were referred after March of 2016 and did not complete treatment. The characteristics of the youth who completed MST and FFT are described in the tables below.

TABLE 10

Gender, JAIS, and Mental Health Disorder, for Completers (2016-2018)

	MST	FFT
GENDER		
Male	42 (74%)	45 (82%)
Female	15 (26%)	10 (18%)
JAIS		
High risk	2 (4%)	10 (18%)
Moderate risk	19 (33%)	26 (47%)
Low risk	18 (32%)	17 (31%)
Missing	18 (32%)	2 (4%)
MENTAL HEALTH DISORDER		
None	5 (9%)	2 (4%)
Externalizing	24 (42%)	20 (37%)
Internalizing	8 (14%)	14 (25%)
PTSD	1 (2%)	8 (15%)
>1	5 (9%)	8 (15%)
Missing	14 (25%)	3 (5%)

TABLE 11

Age of MST and FFT Completers (2016-2018)

Variable	<i>M</i>	<i>SD</i>	<i>n</i>	<i>SE_M</i>
Age				
MST	16.06	1.36	57	0.18
FFT	17.18	1.87	55	0.25

Among the 57 youth who completed MST between March 2016 and February 2018, 24.56% were rearrested for any crime and 5.26% were rearrested for a violent crime. Among the 55 youth who completed FFT during the same time period, 47.27% were rearrested for any crime and 21.82% were rearrested for a violent crime. The rearrest data for completers are depicted in Table 12 and Table 13.

TABLE 12

Number of Charges for first Rearrest for MST and FFT Completers (2016-2018)

	None	1	2	3	4	5	6	7
MST	43(75%)	5 (9%)	5 (9%)	2 (4%)	1 (2%)	1(2%)	0	0
FFT	29 (53%)	13 (24%)	6 (11%)	3 (5%)	1 (2%)	1 (2%)	1 (2%)	1 (2%)

TABLE 13

Charges by Crime Type for first Rearrest for MST and FFT Completers (2016-2018)

Violent	MST	FFT	Drug	MST	FFT
0	54 (95%)	43 (78%)	0	56 (98%)	54 (98%)
1	3 (5%)	11 (20%)	1	1 (2%)	0 (0%)
2	0 (0%)	1 (2%)	2	0 (0%)	1 (2%)
Property	MST	FFT	Gang	MST	FFT
0	50 (88%)	40 (73%)	0	57 (100%)	54 (98%)
1	5 (9%)	12 (22%)	1	0 (0%)	0 (0%)
2	2 (4%)	2 (4%)	2	0 (0%)	1 (2%)
3	0 (0%)	1 (2%)	3	0 (0%)	0 (0%)
Weapons	MST	FFT	Other	MST	FFT
0	54 (95%)	49 (89%)	0	48 (84%)	48 (87%)
1	1 (2%)	3 (5%)	1	7 (12%)	6 (11%)
2	0 (0%)	1 (2%)	2	2 (4%)	1 (2%)
3	2 (4%)	0 (0%)			
4	0 (0%)	1 (2%)			
7	0 (0%)	1 (2%)			

Note. Due to rounding errors, column wise percentages may not equal 100%. Other category is misdemeanors that could not otherwise be classified (false ID, resisting arrest). One MST youth was rearrested for a prostitution-related charge.

Based on the rearrests for MST versus FFT completers over almost a 2 year period of time, we determined the absolute risk reduction, relative risk reduction, and number needed to treat with MST instead of FFT to reduce the likelihood of an arrest at three points in time. Relative risk reduction (RRR) is a measure of how much one treatment (in this case, MST) reduced the risk of an outcome (rearrest) relative to another treatment (in this case, FFT) (Gordis, 2014). Absolute risk

reduction (ARR) reflects the absolute difference in the rate of rearrests among MST youth versus FFT youth. An ARR of 1 suggests there was no effect of MST on rearrests; an ARR above 1 would favor FFT in terms of reducing the number of rearrests. An ARR below 1 favors MST. The number needed to treat (NNT) is the inverse of the ARR. The NNT provides an easily understood outcome of treatment studies, indicating, in this case, how many youth would need to be treated with MST to result in one less rearrest. However, these findings must be interpreted with caution. They do not take into account the baseline risk differences in the two groups and are more useful when there is a true control group, a larger sample, and a longer period of follow-up (Wen, Badgett, & Cornell, 2005). The results of these calculations can be seen in Table 14.

TABLE 14

Rearrests for MST and FFT with Relative Risk Reduction, Absolute Risk Reduction, and NNT

Time	MST Arrests	FFT Arrests	RRR (95% CI)	ARR (95% CI)	NNT
275 days	2/9	3/10	0.26 (-2.48-0.84)	0.08 (-0.30-0.42)	13
550 days	5/29	6/15	0.71 (0.29-0.88)	0.43 (0.13-0.65)	2
730 days	7/19	17/30	0.35 (-0.27-0.67)	0.20 (-0.08-0.44)	5

Note: Arrests are recorded as those arrested/number of all completers under observation during time period

FFT Participants

The final sample of FFT participants were 315 youth referred to TOSO between October 1, 2015 and June 22, 2018. As expected, the majority were Black and Latino young men with an average age of 17 years residing in the eastern region of Contra Costa County (see Tables below). A large percentage of youth were referred from two zip codes in Antioch and Pittsburg. The majority were of moderate risk (43%) and were referred from the ranch (45%). Four youth were identified as CSEY, but only one was referred from the community partner case managing CSEY. Twelve youth (4%) were referred from the community and were not on probation.

TABLE 15

Age of TOSO Youth

	<i>M</i>	<i>SD</i>	<i>N</i>	<i>SE_M</i>
Age	17.20	1.41	315	0.08

TABLE 16

Gender of TOSO Youth

	Gender
Males	248 (78%)
Females	69 (22%)

TABLE 17

Racial/Ethnic Group of TOSO Youth

	Race/Ethnicity
Black	121 (38%)
White	52 (17%)
Latino	127 (40%)
Asian/Pacific Islander	7 (2%)
Mixed	1 (0.3%)
Missing	5 (2%)

TABLE 18

Area of County of TOSO Youth

	Area of County
West	114 (36%)
Central	35 (11%)
East	161 (51%)
Out of County	4 (1%)
Missing	1 (0.03%)

The number of youth referred to TOSO who had MediCal was lower than expected (60%); 7% also had missing data. After referral for services, youth spent an average of 33 days on a waitlist ($SD=46.79$, $SE_M=2.88$, Range from 0-359). Due to several youth returning to detention after being referred and, therefore, waiting up to a year for FFT services, the median number of days on the

waitlist, 11 days, is a more accurate depiction of time between referral and start of treatment. Of those referred for services, 51 (16%) were never enrolled in FFT. The lowest enrollment was during the second half of the first year of TOSO implementation when all but one of the original, trained FFT clinicians remained and the team was being rebuilt. Two hundred seventy-two (86%) of the 315 families referred to TOSO were enrolled, 264 were opened and seen (84%), and 124 (46%) successfully completed FFT. For families who were enrolled, they had an average of 7 sessions ($M=6.27$, $SD=5.73$, $SE_M=0.35$, range 0-23).

FFT Treatment Teams. Given the expected and unexpected challenges faced in the first year of implementation, it is not surprising that the targets for number of youth and families served were not met. We examined the cases referred, number successfully opened, time on waitlist, number of FFT sessions, and completed cases for four FFT teams: 1) those who implemented the program in the first 6 months, 2) the second 6 months of the first year of TOSO when the team was reconfigured with all but one of the original clinicians, 3) the second full year of TOSO when a new clinical supervisor took over, and 4) the third year, from October of 2017 to June of 2018. The period of time at the end of year two, when team 3 was offering FFT, was the most stable period in terms of the FFT clinical and MIOCR grant management personnel (see Table 19).

TABLE 19

Number referred to TOSO, enrolled, completed (% completed after opening), by Treatment Team

	Team 1	Team 2	Team 3	Team 4
	Oct 2015-Mar 2016	Apr 2016-Sept 2016	Oct 2016-Sept 2017	Oct 2017-July 2018
Referred	66	68	100	81
Opened	48	52	86	78
Completed	24 (50%)	36 (69%)	48 (56%)	16 (21%)
Still Open	0	1	0	39 (46%)

The majority of youth referred during early implementation of TOSO was from the community (33%) and ranch (29%). As referral processes improved, the number of youth referred from in-custody treatment programs increased and almost all youth referred were opened. By October of 2017, two years after implementing TOSO, 46% of referrals were from the ranch, 25%, were from

YOTP, and 16% were from GIM. The remaining youth were referred from juvenile hall, placement, or the community.

The number of cases opened increased from 48 in the first 6 months to 52 in the second 6 months. By the end of the second year, TOSO was serving 86 youth and families an average 3.6 months in treatment. Days on waitlist was reduced over time, from 43 days at the end of the first year of TOSO implementation to 14 days during the last year ($M=42.59$, $SD=36.64$ versus $M=14.35$, $SD=13.05$, $p < 0.001$). Length of treatment and number of sessions increased, with an average of 100 days in treatment and 7 sessions for those referred during the first 6 months of TOSO and an average of 119 days in treatment and 8 sessions for those referred during the second full year after TOSO was implemented. Differences in length of treatment or number of sessions could not be explained by MediCal status, $t(250)=-0.27$, $p=0.79$.

FFT Youth Prior Involvement with Juvenile Justice System. We examined the TOSO youths' prior involvement with the justice system to contribute greater understanding of the risk level of the sample. Twelve (4%) of the youth referred from the community were not on probation though one youth had been detained for a day in juvenile hall and was on informal probation at the time of referral. Those youth with sustained petitions (SPs) prior to FFT were an average age of 16 when they received the first SP ($M=15.67$, $SD=1.52$), with a range from 12-20 years of age; their average age at first detention was also 16 years ($M=15.77$, $SD=1.51$) and a range from 12-22 (see Table 20).

TABLE 20
Age of First Sustained Petition and First Detention for TOSO Referrals

	<i>M</i>	<i>SD</i>	<i>n</i>	<i>SE_M</i>
Age of First SP	15.67	1.52	299	0.09
Age of First Detention	15.77	1.51	288	0.09

Among the entire sample, there were an average of 3 SPs prior to referral to TOSO ($M=3.47$, $SD=2.72$); for those who were opened with at least one session of FFT, they had an average 4 SPs prior to referral ($M=3.93$, $SD=2.68$). The difference between the average number of SPs for those referred and opened without a session and those referred and opened with at least one session was

not statistically significant, $t(291)=0.93, p=0.35$. The TOSO youth had an average of 4 prior detentions ($M=4.37, SD=3.89$) with a range from 0-22; similarly, there was no statistically significant difference in number of prior detentions for those who did and those who did not have at least one session of FFT, $t(291)=1.52, p=0.13$.

FFT Youth Completers versus Non-Completers

Prior to determining the impact of TOSO on recidivism, we examined the sample characteristics for comparison between those who completed FFT and those who did not. There were 124 families who completed all phases of FFT treatment and 99 families who began but did not complete FFT. The most common reasons identified for non-completion were refusal and a return to custody (see Table 21).

TABLE 21
Reasons for Non-Completion of FFT

	N (%)
Ran away	17 (17%)
Returned to custody	32 (32%)
Moved	1 (1%)
Refused	32 (32%)
Other	10 (10%)
Missing	7 (7%)

There were no significant differences between completers and non-completers in terms of referral source for ranch, GIM, juvenile hall, and placement (see Table 22); however, YOTP referrals and youth referred from the community had a significantly greater number of non-completers than expected ($\chi^2(1)=4.25, p=0.039$ for YOTP referrals and $\chi^2(1)=13.89, p<0.001$ for community referrals). Of the 12 youth referred who were not on probation, none completed FFT. Two youth participated in FFT and were shot and killed soon after they completed treatment.

TABLE 22

Referral Source, by FFT Completion Status

	Non-Completers	Completers
Ranch	68 (48%)	64 (45%)
YOTP	28 (53%)	19 (36%)
GIM	19 (42%)	20 (44%)
Juvenile Hall	2 (29%)	2 (29%)
CSEY Program	0 (0%)	1 (100%)
Placement	13 (62%)	8 (38%)
Community	20 (44%)	10 (22%)

Note. Due to rounding, percentages may not equal 100%.

In terms of level of risk determined by JAIS category, those youth categorized as high risk were statistically significantly less likely to complete FFT ($\chi^2(1)=19.11, p < 0.001$); completion rates did not differ for moderate and low risk youth (see Table 23). Completers, then, were more likely to be of low to moderate risk, on probation, and referred from the ranch or GIM.

TABLE 23

JAIS Risk Category, by FFT Completion Status

	Non-Completers	Completers
Low Risk	36 (40%)	44 (49%)
Moderate Risk	65 (47%)	62 (45%)
High Risk	38 (67%)	12 (21%)

Note. Due to unknown JAIS status for 31 youth, total sample not included.

We also examined the demographic characteristics of non-completers versus completers. The average age for non-completers and completers was the same (17 years; see Table 24). There were a greater number of both male and female youth among the non-completers than expected; we further explored the differences between male and female youth and found that gender did not have an effect on the odds of completing FFT ($\chi^2(1)=0.23, p=0.629$, McFadden $R^2=0.00$).

TABLE 24

Age of Youth, by FFT Completion Status

	<i>M</i>	<i>SD</i>	<i>n</i>	<i>SE_M</i>
Age of Non-Completers	17.22	1.55	150	0.13
Age of Completers	17.31	1.16	124	0.10

Racial/ethnic group appeared to influence FFT completion, with less Black youth completers than expected, $\chi^2(1)=19.84, p < 0.001$ (see Table 25). Geographic area also influenced completion of FFT; rates of completion did not differ for youth in the west and central areas of the county but there were a greater number of non-completers than would be expected by chance among those in the east area of Contra Costa County ($\chi^2(1)=18.79, p < 0.001$).

TABLE 25

Racial/Ethnic Group, by FFT Completion Status

	Non-Completers	Completers
Black	70 (58%)	37 (31%)
White	22 (42%)	23 (44%)
Latino	50 (39%)	60 (47%)
Asian/Pacific Islander	3 (43%)	3 (43%)
Mixed	1 (100%)	0 (0%)
Other	0 (0%)	1 (100%)
Missing	2 (67%)	0 (0%)

Note. Due to rounding, percentages may not equal 100%.

The results of a chi-square test of independence revealed a statistically significant difference in MediCal status ($\chi^2(1)=7.60, p=0.006$), with a greater number of non-completers than expected among those with MediCal. The odds ratio (OR), for odds of completion among those with MediCal did not provide evidence that the difference amounted to a clinically significant effect. The OR was 0.91 (95% confidence interval, 0.56-1.48). The wider the confidence interval, the less precise the estimated value; when a confidence interval includes 1, as this OR does, there is not likely to be a

statistically significant difference between groups.

Last, we evaluated the effect of prior involvement with the justice system on successful completion of FFT. We used a point biserial correlation, similar to a Pearson’s correlation, to determine the relationship between the number of prior detentions, prior sustained petitions (SPs), prior felony SPs and completion of FFT. We also evaluated the relationship between prior SPs for violent, property, drug, gang, and weapons charges. The number of prior detentions was negatively correlated with FFT completion but the number of pre-FFT SPs was positively correlated with FFT completion. This difference may be due to probation supervision for those with a greater number of prior sustained petitions, but we had no way to test this hypothesis. The effect size for all correlations was small (Cohen, 1988), indicating the relationship between prior involvement with the justice system and FFT completion was not likely to be clinically significant. There were no statistically significant relationships found between number of violent, property, drug, gang, or weapon charges and completion of FFT (see Table 26).

TABLE 26
Correlations between Prior Justice System Involvement and FFT Completion

	<i>n</i>	<i>r_{pb}</i>	<i>95% CI</i>	<i>p</i>
Completion and pre-FFT detentions	274	-0.19	-0.30, -0.07	.002
Completion and age of first detention	256	0.16	0.04, 0.28	.010
Completion and pre-FFT SPs	265	0.18	0.06, 0.29	.004
Completion and age of first SP	274	-0.17	-0.28, -0.05	.006
Completion and number violent offenses	274	-0.05	-0.16, 0.07	.440
Completion and number property offenses	274	-0.04	-0.15, 0.08	.54
Completion and number weapon offenses	274	0.04	-0.08, 0.16	.510

Note. Eighteen youth (0.06%) did not have a SP or detention prior to referral

Youth and Family Functioning before and after Completion of FFT

We examined the effect of FFT on secondary outcomes related to youth and family functioning. Three pre-post measures were collected to assess the family dynamics as they changed as a result of treatment: the outcome questionnaire (OQ), youth outcome questionnaire (YOQ) and the youth outcome questionnaire, self-report (YOQ-SR). We also collected data about pre-post school

attendance and employment. Two hundred thirty-eight (90%) of the 265 families opened in FFT completed an OQ, 233 (88%) completed a YOQ, and 245 (92%) completed a YOQ-SR during the first session of treatment. We examined descriptive statistics about the total and subscale scores for these assessments to provide information about youth and family functioning at the beginning of FFT treatment. The pre-OQ mean total score was 41.46 with a standard deviation of 22.54, indicating the scores were very spread out (see Table 27). The mean total scores and the subscale scores did not reach clinical cut scores.

TABLE 27
Pre-FFT Outcome Questionnaire Subscale and Total Scores

	<i>M</i>	<i>SD</i>	<i>n</i>	<i>SE_M</i>
Symptom distress	22.40	14.10	228	0.93
Interpersonal relations	11.33	6.53	228	0.43
Social role performance	7.73	4.50	228	0.30
TOTAL	41.46	22.54	228	1.49

The pre-YOQ subscale and total scores were explored for parent-report of youth behavior and functioning prior to family treatment (see Table 28). Similarly, the mean subscale scores and the mean of the total score did not reach the clinical cut points; however, the mean score for parent-report of critical items (measuring suicidal ideation and self-harm, psychosis, severe anxiety, and manic behaviors) was 4.56, approaching the clinical cut score of 5. The range on this one measure among the entire sample was 0-32; one youth was reported to ‘almost always’ have all the symptoms measured by the 9-items on the questionnaire.

TABLE 28

Pre-FFT Youth Outcome Questionnaire Subscale and Total Scores

	<i>M</i>	<i>SD</i>	<i>n</i>	<i>SE_M</i>
Interpersonal distress	13.10	10.69	222	0.72
Somatic symptoms	3.90	4.48	223	0.30
Interpersonal relations	5.28	7.15	222	0.48
Social problems	5.72	6.06	222	0.41
Behavioral dysfunction	9.69	8.13	223	0.54
Critical items	4.56	4.85	223	0.32
TOTAL	41.03	35.71	229	2.36

Pre-YOQ-SR total and subscale score means were similar to YOQ scores; the mean subscale score measuring critical items by youth self-report (5.46) was above the clinical cut score of 5 (see Table 29). The mean subscale score for somatic symptoms (4.97) was approaching the clinical cut score for this subscale, 5; the mean subscale scores for interpersonal relations (4.91) and social problems (5.12) were above the clinical cut scores of 4.4 and 3, respectively. These results indicate that FFT youth were, on average, reporting psychiatric and social problems similar to youth reports among those with mental health disorders.

TABLE 29

Pre-FFT Youth Outcome Questionnaire-Self Report Subscale and Total Scores

	<i>M</i>	<i>SD</i>	<i>n</i>	<i>SE_M</i>
Interpersonal distress	14.45	11.18	234	0.73
Somatic symptoms	4.97	4.21	234	0.28
Interpersonal relations	4.91	5.93	234	0.39
Social problems	5.12	4.82	234	0.32
Behavioral dysfunction	10.45	7.49	234	0.49
Critical items	5.46	5.02	234	0.33
TOTAL	45.35	32.31	234	2.11

We calculated Cronbach alpha coefficients to examine overall reliability of the pre-FFT OQ, YOQ, and YOQ-SR measures and found good internal consistency (α s of 0.83, 0.80, and 0.80). We also examined Pearson product-moment correlations of the YOQ and YOQ-SR total scores and correlations of the subscales with average scores greater than the clinical cut scores (see Table 30). Significant positive correlations were observed between all subscale scores and total scores except for the non-significant positive correlation of YOQ social problems and YOQ-SR somatic symptoms. There were significant and moderate correlations between the parent-reported YOQ and the youth reported YOQ-SR subscale scores (0.34 for somatic symptoms, 0.40 for interpersonal relations, 0.45 for social problems, and 0.36 for the critical items). The correlation of YOQ and YOQ-SR total scores was also moderate (Cohen, 1998). The strongest correlation coefficient was with the critical items subscale and the total score. These results suggest that parent and youth reports of behavior, social relations, and family functioning were positively and directly correlated.

TABLE 30

Pearson Product-Moment Correlations of Pre-FFT YOQ and YOQ-SR

	1	2	3	4	5	6	7	8	9	10
1. Somatic symptoms (YOQ)	-									
2. Interpersonal relations (YOQ)	0.56	-								
3. Social problems (YOQ)	0.62	0.76	-							
4. Critical items (YOQ)	0.68	0.67	0.68	-						
5. YOQ total score (YOQ)	0.74	0.89	0.85	0.83	-					
6. Somatic symptoms (YOQ-SR)	0.34	0.17	0.16	0.20	0.27	-				
7. Interpersonal relations (YOQ-SR)	0.29	0.40	0.39	0.33	0.44	0.56	-			
8. Social problems (YOQ-SR)	0.30	0.31	0.45	0.32	0.39	0.44	0.66	-		
9. Critical items (YOQ-SR)	0.30	0.23	0.24	0.36	0.34	0.67	0.62	0.54	-	
10. YOQ-SR total score	0.39	0.37	0.36	0.39	0.47	0.78	0.81	0.69	0.85	-

All correlations were significant at $p < .05$ except that of YOQ SP and YOQ-SR S

We examined descriptive statistics about the total and subscale scores for the OQ, YOQ, and YOQ-SR assessments to provide information about youth and family functioning at the end of FFT treatment before comparing the pre-post scores. One hundred seventeen (94%) of the parents of the 124 families who completed FFT submitted the OQ, 122 (98%) submitted the YOQ and the YOQ-SR (for 2 parents and 2 youth, we had only total score). The mean subscale and total scores for the post-FFT OQs were all in the non-clinical range (see Table 31).

TABLE 31

Post-FFT Outcome Questionnaire Subscale and Total Scores

	<i>M</i>	<i>SD</i>	<i>n</i>	<i>SE_M</i>
Symptom distress	19.68	14.62	117	1.35
Interpersonal relations	10.75	6.42	117	0.60
Social role performance	7.37	4.67	117	0.43
TOTAL	37.70	22.96	117	2.12

The post-FFT total and subscale mean scores were also examined for any that met the criteria for cut scores, indicating clinically significant scores (see Table 32). All mean scores were lower post-FFT than pre-FFT and none approached the clinical cut points. Though the post-score was lower than the pre-score on the YOQ, there remained a lot of dispersion around the mean.

TABLE 32

Post-FFT Youth Outcome Questionnaire Subscale and Total Scores

	<i>M</i>	<i>SD</i>	<i>n</i>	<i>SE_M</i>
Interpersonal distress	10.48	9.59	120	0.88
Somatic symptoms	3.14	3.52	120	0.32
Interpersonal relations	3.62	5.89	120	0.54
Social problems	3.47	4.46	120	0.41
Behavioral dysfunction	6.64	7.44	120	0.68
Critical items	3.10	3.04	120	0.28
TOTAL	29.96	29.25	122	2.65

Post-FFT YOQ-SR mean subscale and total scores were also lower following treatment (see Table 33). No mean scores were above the clinical cut points.

TABLE 33

Post-FFT Youth Outcome Questionnaire-Self Report Subscale and Total Scores

	<i>M</i>	<i>SD</i>	<i>n</i>	<i>SE_M</i>
Interpersonal distress	11.59	10.68	122	0.97
Somatic symptoms	4.34	4.50	122	0.41
Interpersonal relations	3.39	5.59	122	0.51
Social problems	2.70	4.08	121	0.37
Behavioral dysfunction	8.04	7.05	122	0.64
Critical items	4.43	5.65	122	0.51
TOTAL	33.91	31.30	124	2.81

Cronbach alpha coefficients for the three post-FFT measures were the same as for the pre-FFT measures; 0.82 for the OQ, 0.80 for the YOQ, and 0.80 for the YOQ-SR, indicating good internal consistency and test-retest reliability (Koo & Li, 2016). We evaluated the pre-post OQ, YOQ, and YOQ-SR data to determine whether they met the assumptions of normality and homogeneity of variance and then conducted paired t-tests to examine pre-post youth and family functioning.

Pre-post OQ. The result of the paired samples t-test for total OQ score was significant, $t(112)=2.41$, $p=0.018$, suggesting that the true difference in the means of pre- and post-OQ was significantly different from zero. Cohen's d , a measure of effect size, was 0.16, indicating the pre-post reduction in OQ among parents participating in FFT was small. The mean pre-OQ score of 41.01 was significantly higher than the mean post-OQ score of 37.21. The findings indicate that parents who participated in FFT had an average 4 point reduction in total OQ score after treatment. A 15 point change in score is an indication of improvement following treatment (Wells et al., 1996); participation in FFT did not effectively reduce total OQ scores among the parents in this sample.

Pre-post YOQ. A paired samples Wilcoxon signed rank test was conducted to examine whether the difference between pre-FFT and post-FFT YOQ total scores was significantly different from zero. The Wilcoxon signed rank test is a non-parametric alternative to the paired samples t-test, used when the normality and homogeneity of variance assumptions required for the t-test have been violated (McDonald, 2014). The results of the Wilcoxon signed rank test were

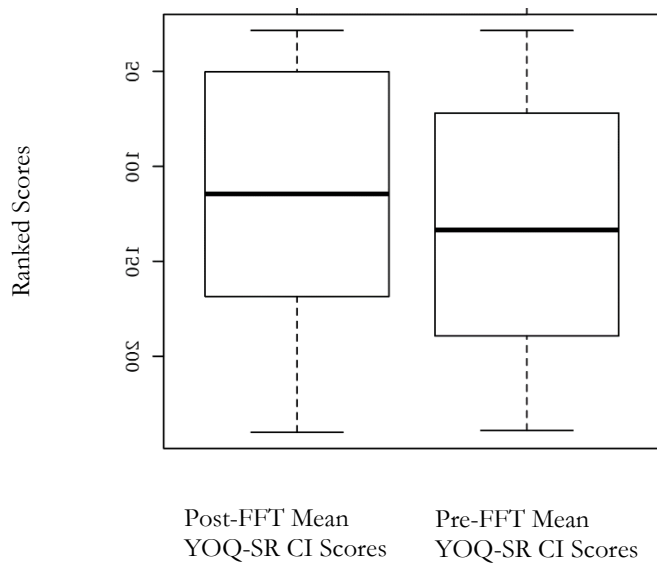
significant, $V=5270.50$, $z=-4.73$, $p<0.001$. This indicates that the differences between the pre- and post-YOQ scores were not likely due to chance. The mean pre-YOQ score of 41.79 was significantly higher than the mean post-YOQ score of 29.54; the change in mean scores as a result of participation in FFT averaged 12 points, one point less than the 13 point reduction used as the indicator of improvement during or following treatment (Dunn et al., 2005).

Pre-post YOQ-SR. The result of a paired samples t -test was significant, $t(121)=2.97$, $p=0.004$ for pre-FFT YOQ-SR total score and post-FFT YOQ-SR total score. The mean of the pre-YOQ-SR total score was 41.39, significantly higher than the mean of the post-YOQ-SR total score, 34.05. The 7 point reduction in the average pre-post YOQ-SR scores suggests that participation in FFT was associated with improved youth behavior and family functioning. However, the mean difference in pre-post scores in this sample did not meet the 13 point reduction that Dunn et al. (2005) reported as an expected index of clinical change.

The strongest correlation on the YOQ and YOQ-SR scores was the subscale score for the critical items and the total score. The average pre-FFT YOQ-SR score for the critical items subscale was above the clinical cut point. For these reasons, we also conducted a Wilcoxon signed rank test to examine whether there was a significant difference between pre-post mean scores for the critical items. The results of the Wilcoxon signed rank test were significant, $V=3100.50$, $z=-2.19$, $p=0.028$, indicating that the difference between the pre-post mean scores was not likely due to random variation (see Figure 3). The mean pre-FFT score for the critical items, 5.07, was above the clinical cut score (5) and the mean post-FFT score for the critical items, 4.4, was below the clinical cut score. The mean difference was 0.66 (95% confidence interval of -0.31 to 1.63). Cohen's d was 0.12, suggesting that completion of FFT contributed a small but significant effect on symptoms of severe mental illness such as suicidal ideation and psychosis among the youth in this study.

FIGURE 3

Box Plot of ranked values for Pre-Post YOQ-SR Scores for the Critical Items (CI)



Last, we examined information about pre-post school attendance and employment. One hundred sixty-seven families reported information about school attendance prior to the start of FFT and 111 reported information at completion. Fifty-nine youth (19%) had graduated from high school or obtained a General Education Diploma (GED) when they enrolled in FFT and 8 were enrolled in community college. One hundred twenty-six youth (42%) were attending school and 106 (35%) had attended more than 20 days in the prior month; however, data were missing about number of days attended for the remainder of the youth. The mean number of school days attended in the previous month was 13 ($SD=9.04$). At point of enrollment in TOSO, 25 (8%) of youth were employed.

Of those families who reported post-FFT information about the youth's education and employment status, 18 (6%) graduated from high school or obtained a GED and 12 youth (4%) were enrolled in community college. Nineteen (16%) FFT completers were not attending school but 71 (58%) were attending and the mean number of school days attended in the previous month increased to 14 ($SD=8.80$). Additionally, the number of youth who were employed increased to 34 (27%). Sixteen youth were both working and going to school. Eighty-nine youth who completed FFT were either attending school, working, or both; 72% of all FFT completers.

While an increase in employment is to be expected, considering that many of the youth had been in-

custody prior to TOSO enrollment and unable to work, the increase in mean number of school days attended is, if not statistically significant, clinically significant. Maintaining progress in school in the community after a return from custody is a challenge for a majority of youth; in fact, results of a national survey suggest about two-thirds do not continue to attend school after being in custody (Federal Interagency Reentry Council, 2012).

FFT Youth Recidivism

Adherence to court-ordered probation. We examined all FFT youth's adherence to court-ordered probation, measured as an absence of any technical violation of probation submitted to the juvenile court. There were 275 post-TOSO enrollment violations of probation among all FFT youth. These violations were committed, however, by only 84 (27%) of the 315 youth, indicating the data were not normally distributed. The average number of violations among the 84 youth was 0.52 ($SD=0.98$, range from 0-7). There were a greater number of violations of probation among FFT non-completers ($M=0.58$, $SD=0.48$, $n=143$) than completers ($M=0.48$, $SD=1.0$, $n=124$).

Using a Kruskal-Wallis rank sum non-parametric test, we assessed if there were significant differences in number of probation violations for non-completers versus completers; results of the Kruskal-Wallis test were not significant, $\chi^2(2)=4.97$, $p=0.08$. A Spearman correlation analysis was conducted between number of probation violations and number of sustained petitions for felony offenses; significant positive correlation was observed ($r_s=0.13$, $p=0.027$) but the effect size was small (Cohen, 1998). Results of these analyses suggest that FFT completion did not improve adherence to court-ordered probation but adherence was associated with a reduction in sustained petitions for felony offenses.

Post-FFT Sustained Petitions for Non-Completers versus Completers. We calculated descriptive statistics to determine the percent of FFT completers who had sustained petitions (SPs) and the frequency of violent, property, drug, gang, and weapon charges post treatment. The 12 youth who were not on probation had to be removed from an analysis of recidivism since they had no probation or court records for review. Sadly, there were also two youth who were killed after completing FFT and they too were removed from the analysis. After removing those youth still receiving FFT, those on probation, and the two who were killed, we were left with data about 215

youth, 91 (42%) non-completers and 124 (58%) completers.

We examined descriptive statistics for sustained petitions by type of crime for the entire sample and then for non-completers versus completers. Among the entire sample, the majority occurrence after participation in TOSO was a violation of probation, with an average of 0.56 ($SD= 1.02$) petitions for violations. All data for sustained petitions were highly skewed, limiting our ability to make inferences about return to crime among youthful offenders and participation in FFT.

Property charges, as an example, which accounted for the most frequent crime type associated with a sustained petition, averaged 0.06 for all youth with SPs (see Table 34). When the skewness is greater than 2 in absolute value, the variable is considered to be asymmetrical about its mean. When the kurtosis is greater than or equal to 3, then the variable's distribution is markedly different than a normal distribution in its tendency to produce outliers (Westfall & Henning, 2013).

TABLE 34

Sustained Petitions for Property Offenses 6-18 months after Exit from TOSO

	<i>M</i>	<i>SD</i>	<i>n</i>	<i>SE_M</i>	<i>Skewness</i>	<i>Kurtosis</i>
SPs for property offenses	0.06	0.26	215	0.02	4.26	19.01

The distribution of property offenses post-FFT is one example of how non-normal the sustained petition (SP) distributions were. Fortunately, there were few post-FFT SPs and the crimes were committed by a small number of youth. Among all observations of SPs following enrollment in TOSO, 22 youth (18%) had a SP for violent, property, or weapon charges, equally distributed between those who did (13 youth, 10%) and those who did not (9 youth, 10%) complete FFT.

For all sustained petitions post TOSO, the mean number of felony charges was greater ($M=0.12$, $SD=0.44$, $SE_M=0.03$, Range=0-4) than the mean number of misdemeanor charges ($M=0.05$, $SD=0.25$, $SE_M=0.02$, Range=0-2). However, these data were also not normally distributed around the mean, so the median provides a more accurate representation of the data. The median number of sustained petitions for felony, misdemeanor, and probation violations was 0. Among the entire sample then, sustained petitions between 6-18 months after completion or non-completion of FFT

were rare.

To test the possibility of a dose-response relationship between participation in FFT and the outcome of any SP for a serious offense post-treatment, we dichotomized the number of FFT sessions into two levels: those who achieved 6 sessions (and likely also achieved the behavior change phase of FFT) and those did not achieve 6 sessions. There were 137 families with 6 or more sessions and 78 with less than 6 sessions. The non-parametric Mann-Whitney two-sample rank-sum test was used to examine whether there were significant differences in sustained petition for a felony offense and FFT dose. The result was not significant, $U=5384.5$, $z=-0.20$, $p=0.844$. The mean rank for group 6 or less was 108.53 and the mean rank for group over 6 was 107.7. This suggests that the distribution of SPs for a serious offense was not different for youth and families who achieved 6 sessions and those who did not achieve 6 sessions of FFT.

We also examined pre-post sustained petitions (SPs) and detentions for youth who did and did not complete FFT. The mean number of SPs for non-completers was greater than for completers before TOSO enrollment (see Table 35). However, the mean number of SPs post-FFT were greater among the entire sample of completers versus non-completers. The range for number of SPs among non-completers, pre-FFT was 0-14 and among completers, pre-FFT was 0-12. The range for number of SPs post-FFT was 0-5 for non-completers and 0-7 for completers. The post-FFT SP data were not normally distributed.

TABLE 35
Pre-Post Sustained Petitions, by FFT Completion

	<i>M</i>	<i>SD</i>	<i>n</i>	<i>SE_M</i>
Pre-FFT SPs, non-completers	4.25	2.98	91	0.31
Post-FFT SPs, non-completers	0.85	1.23	91	0.13
Pre-FFT SPs, completers	3.06	2.15	124	0.19
Post-FFT SPs, completers	0.63	1.21	124	0.11

We used the Wilcoxon Signed Rank test to examine whether there was a statistically significant difference between pre-post SPs for youth who did and did not complete FFT. The results of the

Wilcoxon signed rank test for non-completers were significant, $V=3539.00$, $z=-7.52$, $p<0.001$. This indicates that the differences between the number of pre-SPs and the number of post-SPs are not likely due to random variation. The results of the Wilcoxon signed rank test were also significant for completers, $V=6384.00$, $z=-8.56$, $p<0.001$. Though we cannot determine the effect of FFT completion on the reduction in SPs post-TOSO, it may be that any participation in FFT was protective. Pre-FFT SPs, however, were collected retrospectively over the lifetime of residencies in Contra Costa County. Post-FFT SPs were collected for only 6-18 months post treatment.

FFT non-completers had an average of 5.58 detentions prior to enrollment in TOSO ($SD=4.33$, $SE_M=0.45$, Range 0-18). Completers had an average 3.76 number of detentions prior to FFT ($SD=3.17$, $SE_M=0.28$, Range 0-22). In contrast to post SPs among non-completers and completers, the average number of detentions post-FFT for non-completers was higher ($M=1.31$, $SD=1.91$, $SE_M=0.20$) than among completers ($M=0.64$, $SD=1.52$, $SE_M=0.14$). We used a Wilcoxon signed rank test to examine whether there was a significant difference between pre-post FFT number of detentions for non-completers; the results were significant, $V=2907.5$, $z=-6.82$, $p<0.001$. There was also a significant difference in pre-post FFT number of detentions for completers, $V=6482$, $z=-8.55$, $p<0.001$. These results suggest that completion of FFT was associated with a significant reduction in detentions though, again, time under observation was not accounted for in this analysis.

As serious offenses post-TOSO were committed by just 22 youth, we focused on more fully describing the characteristics of these youth and their families. Two of the youth were female and referred from the GIM program (one completed FFT and one did not); the remaining youth were male (12 completed FFT and 8 did not). One male youth was referred from placement and another from the community; the remaining youth were all referred from the ranch. Treatment completers had an average of 13 sessions of FFT ($SD=3.88$) and non-completers had an average of 4 sessions of FFT ($SD=2.89$). There were no significant differences between non-completers and completers in terms of race/ethnicity. For the non-completers, 3 were Black, 2 were White, and 4 were Latino. Among the completers, 4 were Black, 3 were White, and 6 were Latino.

The residence zip codes for 75% the youth who did not complete FFT and had a sustained petition after referral were located in East County; specifically, in Oakley, Pittsburg, and Bay Point. The rest

of the non-completers lived in North Richmond. Three of the youth who had a sustained petition and completed FFT lived in Pittsburg; overall, 50% with a sustained petition post-FFT were from West County, 33% from East County, and 17% from Central County (see Appendix D).

Comparing the two female youth, age and prior involvement with probation likely played a role in treatment success or failure. The female who did not complete treatment had her first sustained petition and her only detention at age 17; the female who completed FFT had first been arrested and detained at age 12, had 9 prior out-of-home placements, and had been identified as commercially sexually exploited. Pre-FFT, both girls had sustained petitions for violent felony charges. Post-FFT, the non-completer had a sustained petition for a violent felony charge 2 months after she exited TOSO and the completer had a sustained petition for two misdemeanor property crimes 14 months after completing FFT.

For the males who had a sustained petition after exit from TOSO, they had an average age of 16.5 (Range 14.8-18.5) at the time of referral. Those who completed were equally divided by MediCal status (42% with, 58% without) but 88% of non-completers with a sustained petition had MediCal. None of the youth were employed and less than half were attending school. Type of crime for the youth who did not complete FFT were primarily property charges (see Table 36). The overall number of felony convictions for non-completers and completers was 13 and 12, respectively.

TABLE 36

Number of TOSO Youth with a Sustained Petition for a Serious Charge, by Completion Status

	N Non-Completers	N Completers
Violent charge	2	7
Property charge	7	6
Weapon charge	2	2

Note: One youth had a sustained petition for attempted murder, another for a violent charge with gang enhancement.

Prior to our review of sustained petitions, we had collected data about rearrests among 33 youth who completed FFT. Twelve youth who completed FFT were identified as having an arrest for a violent crime; as of the date we collected information about sustained petitions, seven of these youth

had a sustained petition. Twelve youth had been arrested for property crimes; half of these arrests resulted in a sustained petition. Only two of the six youth arrested and charged with a weapon crime had the petition sustained. Overall, the majority of the 22 probation youth served by TOSO who had a sustained petition post-referral were low-income (MediCal eligible) males from East County, of moderate risk, referred after being remanded to the ranch.

TOSO Program Success

The following were set as targets that, if met, would provide evidence of the effectiveness of TOSO: 1) 20% reductions in re-arrests among YOTP, GIM and ranch participants over the years 2015-2018, 2) not more than 10% of GIM participants returning to the GIM program by 2018, 3) completion of FFT by more than 70% of families referred, 4) maintenance of the youth in the home with significant improvements in pre-post YOQs for those completing FFT and 4) 75% of youth in families participating in FFT attending school and/or employment at completion of program.

TOSO achieved almost all targets, as follows:

1) Youth exiting YOTP from 2013-2015 had a rearrest rate of 35% 6-9 months post-release (Gerchow, 2015). Among this sample, no YOTP youth had a sustained petition 6-18 months post treatment. As of 2016, 48% of GIM participants had been rearrested but they were followed anywhere between 6 months to 6 years (Shade & Ceric, 2017). Among this sample, 1 (0.8%) GIM youth had a sustained petition 6-18 months post treatment. We do not have recidivism data for the ranch; in this sample, youth referred from the ranch were most likely than youth referred from other sources to have a sustained petition 6-18 months post treatment.

2) One GIM youth who had been commercially sexually exploited returned to custody after completion of TOSO.

3) Of the 264 families opened in TOSO from 2016-2018, 124 (47%) completed and 40 are on track to complete. If they do, TOSO will be 10% below the 70% target for completion for referred families.

4) There were significant improvements in youth behavior, psychological symptoms, and family functioning as measured by YOQs for those families who completed TOSO.

5) At time of exit from TOSO, 72% of youth were either attending school, employed, or both.

COSTS FOR TOSO

We collected the costs of the program for the 3-year grant period in order to provide an estimated cost for a county to offer a TOSO program without grant support. The costs for grant management and program evaluation are not included. We collected both MST and FFT costs to Contra Costa County in the form of payments to the community-based contract provider, COFY, and required training costs. We also collected the total costs, in salaries and benefits, paid by the county behavioral health department to support the TOSO project (See Table 37). Costs were returned to the county in the form of the federal match for MediCal services. Sixty-five percent of MST youth and 68% of FFT youth were covered by MediCal.

TABLE 37

Cost to Deliver MST and FFT to Probation Youth in Contra Costa County (Oct 2015-June 2018)

	MST	FFT
Direct services	\$2,476,434	\$1,477,839
Lockout Services (grant-funded)	\$88,738	\$98,243
Training and Consultation	\$220,176	\$27,121
Behavioral Health Team Services	\$269,204	\$122,308
TOTAL	\$3,054,552	\$1,725,511

Note: Costs are rounded. Lockout services include visits with youth and family while youth in custody; the Early and Periodic Screening, Diagnostic and Treatment (EPSDT) program does not cover the cost of assessment and referral when youth are in custody. Initial training and consultation costs were totaled and averaged over 2.9 years. Behavioral health team included program manager and 3 liaisons; manager provided 10% time to MST and FFT, liaisons provided 30% time to MST and 10-20% time to FFT.

The cost of this 3-year program evaluation totaled \$50,000.

CONCLUSIONS AND RECOMMENDATIONS

Training and support for providers of FFT services. The authors of a systematic review about implementation and sustainability of new programs suggest that the majority of interventions are successful when they are well-supported during the first two years (Wiltsey-Stirman, Kimberly, Cook, Calloway, Castro, & Charns, 2012). As we learned through the process evaluation, capacity, fit, and the effectiveness of a program are important but attention to processes and interactions was critical to the success of TOSO. It took a few false starts before the right team and the right supervisor was configured to implement FFT. As described by Smyth and Schorr (2009), successful social interventions are ones that emphasize relationships and trust, establish working partnerships with participants, account for the larger context in which the program is located, and focus on realistic goals for each individual and family rather than aiming for narrow outcomes. An evidence-based social intervention is effective not because of the intervention itself but because of the effectiveness of the people providing the intervention.

Addressing primary and secondary trauma. We also learned through the process evaluation how important it is to address primary trauma among the youth and families served by TOSO and to prevent, recognize, and treat secondary trauma, or compassion fatigue, among FFT clinicians as well (Pearlman & MacIan, 1995). A concise tool is available to assess for compassion fatigue, the Professional Quality of Life Scale, known as the ProQOL, and could be used to regularly check in with the probation and mental health staff who work on the front-lines with youth and families (Stamm, 2010). A joint training would effectively strengthen and support the partnerships already developed between probation, mental health, and COFY clinicians and would be an opportunity to share information about self-care strategies.

Assessment and enrollment of youth while still in custody. The JAIS did not effectively predict the youth in this study who were rearrested, reconvicted, or returned to custody. As of 2015, GIM probation staff have used the Ohio Youth Assessment System's (OYAS) re-entry risk tool for case planning. The OYAS disposition and the OYAS detention have been found to have adequate predictive validity for both male and female youth, but the other OYAS risk assessment instruments (for diversion and residential case planning) are not as reliable and have not been tested with females. For this reason, field probation officers should be trained in the use of the OYAS

disposition tool and it should be used for placement into the in-custody treatment programs instead of the JAIS (Lovins & Latessa, 2013).

We found through the pre-post evaluation of youth behavior and psychiatric symptoms that the majority self-reported symptoms of serious mental illness. The youth treated through FFT, rather than MST, were also more likely to be suffering from internalizing disorders and posttraumatic stress disorder. Incarcerated youth have much higher rates of difficult-to-treat complex trauma than youth in the general and clinical population (Ford, Chapman, Connor, & Cruise, 2012). In order to effectively intervene at the earliest opportunity, a reliable and valid tool such as the Structured Assessment of Violence Risk in Youth (SAVRY) should be used to identify those at risk who are remanded to in-custody treatment (Borum, Bartel, & Forth, 2006). According to Viljoen, Shaffer, Gray, and Douglas (2017) an additional advantage of the SAVRY is that it is a sensitive measure that can be used to assess change over time during treatment.

Improving transitional services at Orin Allen Rehabilitation Facility (the ranch). The FFT impact evaluation suggested that the majority of youth referred to TOSO and the majority reconvicted and returned to custody after TOSO were ranch youth. The TOSO mental health liaisons and FFT clinicians recognized early in the implementation phase of the program that partnerships with ranch probation officers was a weak link in the referral process. The ranch managers have recently implemented the core correctional practice trainings that have already been instituted in the juvenile hall (University of Cincinnati Corrections Institute; UCCI, n.d.). The practices provide probation personnel with the skills needed to work within the framework of a cognitive-behavioral paradigm, consistent with best practice in managing problem behavior among adolescents and young adults (Landenberger & Lipsey, 2005).

As the ranch team learns more about cognitive-behavioral approaches, we recommend they also increase the dose of CBT for ranch youth. In 2016, we completed an evaluation of the Board of State and Community Corrections (BSCC) Evidence Based Practice to Improve Public Safety (EBP-TIPS) grant-funded project (Shade & Ceric, 2016). The goal was to improve the fidelity of the Thinking for a Change (T4C) CBT intervention and to increase the number of youth receiving T4C (Bush, Glick & Taymans, 1997). We evaluated pre-post improvement in cognitive distortions and self-reported criminal behavior using the 'How I Think' (HIT) questionnaire (Gibbs, Barriga, &

Potter, 2001). HIT scores correlated with JAIS category. We found more high risk participants were located at GIM and the ranch (average pre-HIT scores of 3.06) than YOTP (average pre-HIT scores of 3.04). Over the course of the quality improvement project, we conducted pre-post assessments on 165 youth (63%) who completed T4C and found 97 youth (37%) started but did not complete the intervention. Ranch, male, and Latino participants were more likely than other youth to be non-completers. The results of our analyses for this project suggest that ranch youth are still at high risk for recidivism and need a higher dose of CBT while in-custody. The FFT clinicians have participated in joint CBT trainings with probation staff and are well-versed at continuing to support youth to ‘think, for a change’ as they transition back to the community.

Supporting youth during the transition to adulthood. MST is delivered to 12-17 years olds based on a developmental view of the needs of adolescents in the context of family, school, and other community systems (Henggeler, 1997). Recent studies suggest that older youth are greatly disadvantaged as a result of aging in the juvenile justice system; their opportunity for positive youth connections and social support deteriorates the longer they remain in-custody (Leverso, Bielby, & Hoelter, 2015; Pettus-Davis, Doherty, Veeh & Drymon, 2017). Jeffery Arnett wrote a seminal work (Arnett, 2000) about the concept of emerging adulthood, identifying a new post-adolescent period that has arisen as a result of economic and social changes and an improved understanding of cognitive development among 18-26 year olds. Arnett argues that emerging young adults require significant support during this transition in order to avoid an array of problems that can have consequences over the course of a lifetime. The challenge of navigating emerging adulthood has been recognized for former foster youth (Greeson, 2013) and racial-minority, immigrant youth (Syed & Mitchell, 213) but has not been explored as thoroughly for juveniles in the justice system.

Recently, developers of MST (Sheidow, McCart, & Davis, 216) have addressed the needs of emerging adults with severe mental illness and justice-involvement. They piloted a project called Multisystemic Therapy for Emerging Adults (MST-EA), designed for youth aged 17-21 reentering the community after in-custody treatment. The goals of MST-EA are to shift responsibilities from parents to young adults, increase social safety nets, treat trauma-associated conditions, reduce antisocial behaviors and exposure to antisocial peers, and ensure the young adult is receiving effective mental health services. For the participants of the first published study, 76% had no rearrest during treatment and pre-post psychiatric admissions were significantly reduced. As MST is

appropriate for youth displaying behavior problems, while FFT is appropriate for youth who are experiencing symptoms associated with trauma (Baglivio et al., 2014) and internalizing disorders, it would be ideal to fund an arsenal of effective family service interventions for Contra Costa County probationers. With appropriate risk stratification of families, MST or FFT could be used as a diversion program for families with younger children and as transitional services for older youth. Based on the youth's mental health problems and criminogenic factors (Andrews et al., 2006), either MST or FFT would be delivered to families with 12-21 year olds.

Targeting chronic offenders. It may be that the 22 youth who recidivated following enrollment in TOSO belong to a small group of serious, chronic offenders, known as the 6% (Wolfgang, Figlio & Sellin, 1972) or those with life-course-persistent antisocial behavior (Moffitt, 1993). It may also be that a family therapy intervention cannot adequately address the social, environmental, educational, epigenetic, and economic disparities in which these youth live, work, and play (Braveman, Egerter, & Williams, 2011). A majority of the youth referred to TOSO and recidivating following referral were located in Antioch and Pittsburg in East County or North Richmond in West County (see Appendix D). These are two areas of the county in which the lowest performing schools in the state can be found (Noguci, 2017) and, according to Scorecard (n.d.) vulnerable youth are exposed to the highest levels of air, water, and ground pollutants than any other area in the county. Functional Family Therapy clinicians partnered with parents to ensure TOSO youth were well-connected to any and all resources at their disposal. However, they cannot make positive social and support connections to youth and families if such connections are not present in their communities.

REFERENCES

- Aalsma, M. C., White, L. M., Lau, K. S., Perkins, A., Monahan, P., & Grisso, T. (2015). Behavioral health care needs, detention-based care, and criminal recidivism at community reentry from juvenile detention: A multisite survival curve analysis. *American Journal of Public Health, 105*(7), 1372-1378. doi: 10.2105/ajph.2014.302529
- Achenbach, T.M. (1991). *Manual for the Child Behavior Checklist/4–18 and 1991 profile*. Burlington, VT: University of Vermont Department of Psychiatry
- Alexander, J. F., Robbins, M. S., & Sexton, T. L. (2000). Family-based interventions with older, at-risk youth: From promise to proof to practice. *The Journal of Primary Prevention, 21*(2), 185-205. doi 10.1023/a:1007031219209
- Allen, D. & Grassel, K. (2017). 2016 *Juvenile justice outcome evaluation report: Examination of youth released from the Division of Juvenile Justice in fiscal year 2011-12*. Center for Evidence-Based Corrections, University of California, Irvine and California Department of Corrections and Rehabilitation
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.
- Andrews, D. A., Bonta, J., & Wormith, J. S. (2006). The recent past and near future of risk and/or need assessment. *Crime and Delinquency, 52*(1), 7-27. <http://dx.doi.org/10.1177/0011128705281756>
- Aos, S., Lieb, R., Mayfield, J., Miller, M., & Pennucci, A. (2004). *Benefits and costs of prevention and early intervention programs for youth*. Olympia, WA: Washington State Institute for Public Policy. Retrieved from <http://www.wsipp.wa.gov/pub.asp?docid04-07-3901>
- Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist, 55*(5), 469-480. doi:10.1037/0003-066x.55.5.469
- Babinski, L. M., Hartsough, C. S., & Lambert, N. M. (2001). A comparison of self-report of criminal

involvement and official arrest records. *Aggressive Behavior*, 27(1), 44-54. doi:
[http://dx.doi.org/10.1002/1098-2337\(20010101/31\)27:1](http://dx.doi.org/10.1002/1098-2337(20010101/31)27:1)

Baglivio, M. T., Jackowski, K., Greenwald, M. A., & Wolff, K. T. (2014). Comparison of multisystemic therapy and functional family therapy effectiveness: A multiyear statewide propensity score matching analysis of juvenile offenders. *Criminal Justice and Behavior*, 41(9), 1033-1056. doi:10.1177/0093854814543272

Baird, C., Healy, T., Johnson, K., Bogie, A., Dankert, E.W. & Scharenbroch, C. (2013, October). A comparison of risk assessment instruments in juvenile justice. *National Council on Crime and Delinquency*. Retrieved from
http://nccdglobal.org/sites/default/files/publication_pdf/nccd_fire_report.pdf

Barnoski, R. (2004). *Outcome evaluation of Washington State's research-based programs for juvenile offenders*. Olympia, WA: Washington State Institute for Public Policy.

Becerra, X. (2018). *Juvenile Justice in California, 2017*. Sacramento, CA: California Department of Justice.

Blueprints for Healthy Youth Development. (2015). *Functional Family Therapy*. Retrieved from
<https://www.blueprintsprograms.org/factsheet/functional-family-therapy-fft>

Borum, R., Bartel, P., & Forth, A. (2006). *Manual for the Structured Assessment of Violence Risk in Youth (SAVRY)*. Odessa, FL: Psychological Assessment Resources.

Braveman, P., Egerter, S., & Williams, D. R. (2011). The social determinants of health: Coming of age. *Annual Review of Public Health*, 32(1), 381-398. doi:10.1146/annurev-publhealth-031210-101218

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.

Burlingame, G.M., Lambert, M.J., Reisinger, C.W. Neff, W.M, & Mosier, J. (1995). Pragmatics of

tracking mental health outcomes in a managed care setting. *The Journal of Mental Health Administration*, 22, 226. <https://doi.org/10.1007/BF02521118>

Bush, J., Glick, B. & Taymans, J. (1997, revised 1998). *Thinking for a Change: Integrated cognitive behavior change program*. Washington, D.C.: National Institute of Corrections, U.S. Department of Justice

California Department of Health Care Services (n.d.) *Medi-Cal state inmate program and Medi-Cal county inmate program*. Retrieved from <http://www.dhcs.ca.gov/provgovpart/Pages/MIP.aspx>

California Department of Justice (n.d.) *Conditions for release of Criminal Offender Record Information*. Retrieved from <https://oag.ca.gov/sites/all/files/agweb/pdfs/Conditions-for-Release-CORI.pdf>

California Institute for Behavioral Health Solutions (n.d.) *OQ measures quick start guide*. Sacramento, CA: Author

California Institute for Behavioral Health Solutions (n.d.) *FFT team certification*. Retrieved from <https://www.cibhs.org/fft-team-certification>

California State Attorney General (2014). *Recidivism in California*. Retrieved from <https://oag.ca.gov/dap/evaluation/recidivism>

Center on Juvenile and Criminal Justice (2018a). *Juvenile justice practices in California by county, calendar year 2016*. [data file]. San Francisco, CA: Author

Center on Juvenile and Criminal Justice (2018b). *California juvenile justice funding in five Bay Area Counties shows opportunity for reinvestment*. Retrieved from http://www.cjcj.org/uploads/cjcj/documents/california_state_juvenile_justice_funding_in_five_bay_area_counties.pdf

Cohen, J. (1988), *Statistical power analysis for the behavioral sciences*, (2nd ed). Hillsdale: Lawrence Erlbaum

Connell, C. M., Steeger, C. M., Schroeder, J. A., Franks, R. P., & Tebes, J. K. (2016). Child and case influences on recidivism in a statewide dissemination of multisystemic therapy for juvenile offenders. *Criminal Justice and Behavior*, *43*(10), 1330-1346. doi 10.1177/0093854816641715

Contra Costa County Health Services (2018). *Contra Costa County Mental Health Services Act: Three year program and expenditure plan update, fiscal year 2018 – 2019*. Retrieved from <https://cchealth.org/mentalhealth/mhsa/pdf/Plan-Update-FY-2018-2019.pdf>

Cook, A. K., & Gordon, J. A. (2012). Get him out of my house: Parental competencies of juvenile probationers. *Youth Violence and Juvenile Justice*, *10*(2), 205-223. doi: 10.1177/1541204011418352

Corman, H., & Mocan, N. (2005). Carrots, sticks, and broken windows. *The Journal of Law and Economics*, *48*(1), 235-266. doi:10.1086/425594

Cottle, C. C., Lee, R. J., & Heilbrun, K. (2001). The prediction of criminal recidivism in juveniles: A meta-analysis. *Criminal Justice and Behavior*, *28*(3), 367-394. doi 10.1177/0093854801028003005

Council of State Governments Justice Center (2014). *Measuring and using juvenile recidivism data to inform policy, practice, and resource allocation*. Retrieved from <http://csgjusticecenter.org/wp-content/uploads/2014/07/Measuring-and-Using-Juvenile-Recidivism-Data-to-Inform-Policy-Practice-and-Resource-Allocation.pdf>

Curtis, N. M., Ronan, K. R., & Borduin, C. M. (2004). Multisystemic treatment: A meta-analysis of outcome studies. *Journal of Family Psychology*, *18*(3), 411-419. doi:10.1037/0893-3200.18.3.411

Dodge, K. A., Coie, J. D., & Lynam, D. (2006). *Aggression and antisocial behavior in youth*. Hoboken, NJ: John Wiley & Sons.

Drummond, M.F., & Jefferson, T.O. (1996). Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *British Medical Journal*, *313*(7052), 275–283. <https://doi.org/10.1136/bmj.313.7052.275>

Drummond, M.F., Sculfer, M., Claxton, K., Stoddard, G.L. & Torrance, G.W. (2015). *Methods for the economic evaluation of health care programmes* (4th ed.). Oxford, UK: Oxford University Press

Dunn, T. W., Burlingame, G. M., Walbridge, M., Smith, J., & Crum, M. J. (2005). Outcome assessment for children and adolescents: Psychometric validation of the Youth Outcome Questionnaire 30.1 (Y-OQ®-30.1). *Clinical Psychology and Psychotherapy*, 12(5), 388-401.
doi:10.1002/cpp.461

Eichler, H. G., Kong, S. X., Gerth, W. C., Mavros, P., & Jonsson, B. (2004). Use of cost-effectiveness analysis in health-care resource allocation decision-making: How are cost-effectiveness thresholds expected to emerge? *Value Health*, 7(5), 518-528. doi:10.1111/j.1524-4733.2004.75003.x

Evers, S., Goossens, M., de Vet, H., van Tulder, M., & Ament, A. (2005). Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. *International Journal of Technology Assessment in Health Care*, 21(2), 240-245.
<https://doi.org/10.1017/S0266462305050324>

Fallon, T., Pumariega, A., Sowers, W., Klaehn, R., Huffine, C., Vaughan, T., . . . Grimes, K. (2006). A level of care instrument for children's systems of care: Construction, reliability and validity. *Journal of Child and Family Studies*, 15(2), 140-152. doi 10.1007/s10826-005-9012-y

Farrington, D. P., Jolliffe, D., Hawkins, J. D., Catalano, R. F., Hill, K. G., & Kosterman, R. (2003). Comparing delinquency careers in court records and self-reports. *Criminology* 41, 933–958.
<https://doi.org/10.1111/j.1745-9125.2003.tb01009.x>

Federal Interagency Reentry Council (2012). *Reentry myth buster: On youth access to education upon reentry*. Retrieved from <http://bit.ly/1sxm157>

Ford, J. D., Chapman, J., Connor, D. F., & Cruise, K. R. (2012). Complex trauma and aggression in secure juvenile justice settings. *Criminal Justice and Behavior*, 39(6), 694-724.
doi:10.1001/archpsyc.61.4.403

Functional Family Therapy, LLC (n.d.). *FFT training*. Retrieved from <https://www.fftlc.com/about-fft-training/implementing-fft.html>

Gerchow, C. E. (2015). *Evaluating the impact of two cognitive behavioral therapy programs on recidivism in chronic juvenile offenders*. (10186620 Ph.D.), University of California, Berkeley, Ann Arbor. ProQuest Dissertations & Theses Global database

Gibbs, J. C., Barriga, A. Q., & Potter, G. (2001). *The How I Think questionnaire*. Champaign, IL: Research Press

Glesmann, C. & Irvine, A. (2014, October). Stanislaus County girls' juvenile justice initiative evaluation report. *National Council on Crime and Delinquency*. Retrieved from https://www.nccdglobal.org/sites/default/files/publication_pdf/stanislaus-gjji-report.pdf

Gordis, L. (2014). *Epidemiology* (5th ed.). Philadelphia, PA: Elsevier Saunders

Gordon, D. A., Graves, K., & Arbuthnot, J. (1995). The effect of Functional Family Therapy for delinquents on adult criminal behavior. *Criminal Justice and Behavior*, 22(1), 60-73. doi: <http://dx.doi.org/10.1177/0093854895022001005>

Greeson, J. K. P. (2013). Foster youth and the transition to adulthood: The theoretical and conceptual basis for natural mentoring. *Emerging Adulthood*, 1(1), 40-51. doi:10.1177/2167696812467780

Hajian-Tilaki, K. (2013). Receiver Operating Characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian Journal of Internal Medicine*, 4(2), 627-635

Harris, K.A. (2016). *Juvenile justice in California, 2015*. Sacramento, CA: California Department of Justice.

Harris, K.A. (2014). *Juvenile justice in California, 2013*. Sacramento, CA: California Department of Justice.

Harris, P.W., Lockwood, B. & Mengers, L. (2009). *Defining and measuring recidivism*. Retrieved from <https://www.ncjrs.gov/App/Publications/abstract.aspx?ID=259797>

Henggeler, S.W. (1997). Treating serious antisocial behavior in youth: The MST approach. *Juvenile Justice Bulletin*. Washington, D.C.: Office of Juvenile Justice and Delinquency Prevention, US Department of Justice.

Henggeler, S. W., & Lee, Terry. (2003). Multisystemic treatment of serious clinical problems. In A. E. Kazdin and J. R. Weisz (eds.), *Evidence-based psychotherapies for children and adolescents* (pp. 301-322). New York: The Guilford Press.

Henggeler, S. W., Melton, G. B., & Smith, L. A. (1992). Family preservation using multisystemic therapy: an effective alternative to incarcerating serious juvenile offenders. *Journal of Consulting and Clinical Psychology*, *60*(6), 953-961. <https://doi.org/10.1037/0022-006X.60.6.953>

Henggeler, S. W., & Sheidow, A. J. (2012). Empirically supported family-based treatments for conduct disorder and delinquency in adolescents. *Journal of Marital and Family Therapy*, *38*(1), 30–58. doi:10.1111/j.1752-0606.2011.00244.x

Hernán, M. A. (2010). The hazards of hazard ratios. *Epidemiology*, *21*(1), 13–15. <http://doi.org/10.1097/EDE.0b013e3181c1ea43>

Heatherington, L., Friedlander, M. L., & Greenberg, L. (2005). Change process research in couple and family therapy: Methodological challenges and opportunities. *Journal of Family Psychology*, *19*(1), 18-27. doi:10.1037/0893-3200.19.1.18

Howell, J.C. & Lipsey, M.W. (2004). A practical approach to linking graduated sanctions with a continuum of effective programs. *Training and Technical Assistance Bulletin*. Juvenile Sanctions Center

Institute for Health Metrics and Evaluation (2018, February 26) *Protocol for the global burden of diseases, injuries, and risk factors study*. Retrieved from http://www.healthdata.org/sites/default/files/files/Projects/GBD/GBD_Protocol.pdf

kidsdata.org (n.d.a) *Juvenile felony arrests, by type of offense, 2013*. Retrieved from <http://www.kidsdata.org/topic/386/juvenilearrests-offense/table#fmt=532&tf=73&ch=373,375,376,374,377&sortColumnId=0&sortType=asc>

kidsdata.org (n.d.b) *Juvenile felony arrests, by type of offense, 2015*. Retrieved from <https://www.kidsdata.org/topic/386/juvenilearrests-offense/table#fmt=532&loc=2&tf=84&ch=373,375,376,374,377&sortColumnId=0&sortType=asc>

Klein, N. C., Alexander, J. F., & Parsons, B. V. (1977). Impact of family systems intervention on recidivism and sibling delinquency: A model of primary prevention and program evaluation. *Journal of Consulting and Clinical Psychology, 45*(3), 469-474. doi:10.1037/0022-006X.45.3.469

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting Intraclass Correlation Coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. <http://doi.org/10.1016/j.jcm.2016.02.012>

Landenberger, N. A., & Lipsey, M. W. (2005). The positive effects of cognitive-behavioral programs for offenders: A meta-analysis of factors associated with effective treatment. *Journal of Experimental Criminology, 1*(4), 451-476. doi: 10.1007/s11292-005-3541-7

Laub, J. H., & Sampson, R. J. (2001). Understanding desistance from crime. *Crime and Justice, 28*, 1-69. <https://doi.org/10.1086/652208>

MST Services. (2009). *Multisystemic therapy organizational manual*. Charleston, SC: Author.

Leverso, J., Bielby, W., & Hoelter, L. F. (2015). Back on the streets: Maturation and risk factors for recidivism among serious juvenile offenders. *Journal of Adolescence, 41*, 67-75. doi:10.1016/j.adolescence.2015.02.008

Lovins, B., & Latessa, E. (2013). Creation and Validation of the Ohio Youth Assessment System (OYAS) and Strategies for Successful Implementation. *Justice Research and Policy, 15*(1), 67-93. doi:10.3818/jrp.15.1.2013.67

McDonald, J.H. (2014). *Handbook of biological statistics* (3rd ed.). Sparky House Publishing: Baltimore, MD

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, 22(3), 276–282

Mallett, C. A., Fukushima, M., Stoddard-Dare, P., & Quinn, L. (2013). Factors related to recidivism for youthful offenders. *Criminal Justice Studies*, 26(1), 84-98. doi:10.1080/1478601X.2012.705539

Massetti, G. M., Vivolo, A. M., Brookmeyer, K., DeGue, S., Holland, K. M., Holt, M. K., & Matjasko, J. L. (2011). Preventing youth violence perpetration among girls. *Journal of Women's Health*, 20(10), 1415-1428. doi: 10.1089/jwh.2011.3057

McClendon, D.T., Warren, J., Green, K.M., Burlingame, G.M., Eggett, D.L., & McClendon, R.J. (2011). Sensitivity to change of youth treatment outcome measures: A comparison of the CBCL, BASC-2, and Y-OQ. *Journal of Clinical Psychology*, 67 1, 111-125. doi 10.1002/jclp.20746

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychology Review*, 100(4), 674-701. doi:http://dx.doi.org/10.1037/0033-295X.100.4.674

Murray, J., & Farrington, D. P. (2010). Risk factors for conduct disorder and delinquency: Key findings from longitudinal studies. *Canadian Journal of Psychiatry*, 55(10), 633-642. doi:10.1177/070674371005501003

National Council on Crime and Delinquency (n.d.). *Juvenile Assessment and Intervention System (JAIS)*. Retrieved from <https://www.nccdglobal.org/assessment/juvenile-assessment-and-intervention-system-jais>

National Institute of Justice (2014). *Recidivism*. Retrieved from <https://www.nij.gov/topics/corrections/recidivism/Pages/welcome.aspx>
Nunnally, J. C. (1978). *Psychometric theory* (2nd ed). New York: McGraw-Hill

Noguci, S. (2017, December 8). California's new report cards: Which districts don't make the grade? *The Mercury News*. Retrieved from <https://www.mercurynews.com/2017/12/08/california-releases-school-report-cards-flags-worst-performers/>

Office of the Governor (2017). *Governor's budget 2017-18: Corrections and rehabilitation*. Retrieved from <http://www.ebudget.ca.gov/2017-18/pdf/GovernorsBudget/5210.pdf>

Pearlman, L.A. & MacIain, P.S. (1995). Vicarious traumatization: An empirical study of the effects of trauma work on trauma therapists. *Professional Psychology: Research and Practice*, 26, 558-565.
<https://doi.org/10.1037/0735-7028.26.6.558>

Pettus-Davis, C., Doherty, E. E., Veeh, C., & Drymon, C. (2017). Deterioration of postincarceration social support for emerging adults. *Criminal Justice and Behavior*, 44(10), 1317-1339.
[doi:10.1177/0093854817721936](https://doi.org/10.1177/0093854817721936)

Pourat, N., Zima, B., Marti, A., & Lee, C. (2017, September). California child mental health performance outcomes system: Recommendation report. UCLA Center for Health Policy Research. Retrieved from http://healthpolicy.ucla.edu/publications/Documents/PDF/2017/California_Child_Mental_Health_Performance_Outcomes_System_Recommendation_Report.pdf

Pumariega, A. J., Millsaps, U., Moser, M., & Wade, P. (2014). Matching intervention to need in juvenile justice: The CASII level of care determination. *Adolescent Psychiatry*, 4(4), 305-313.
<https://doi.org/10.2174/221067660404150115163555>

Reynolds, C.R., & Kamphaus, R.W. (2004). *Behavioral Assessment System for Children, Second Edition (BASC-2)* (2nd ed). Circle Pines, MN: AGS Publishing

Robbins, M. S., Alexander, J. F., Turner, C. W., & Hollimon, A. (2016). Evolution of Functional Family Therapy as an evidence-based practice for adolescents with disruptive behavior problems. *Family Process*, 55(3), 543-557. [doi:10.1111/famp.12230](https://doi.org/10.1111/famp.12230)

Robberstad, B. (2009). QALYs vs DALYs vs LYs gained: What are the differences, and what difference do they make for health care priority setting? *Norsk Epidemiologi*, *15*(2) 183-191.
<https://doi.org/10.5324/nje.v15i2.217>

Rodgers, J., & Nicewander, W. (1988). Thirteen ways to look at the correlation coefficient. *The American Statistician*, *42*(1), 59-66. doi:10.2307/2685263

Syed, M., & Mitchell, L. L. (2013). Race, ethnicity, and emerging adulthood: Retrospect and prospects. *Emerging Adulthood*, *1*(2), 83-95. doi:10.1177/2167696813480503

Schaeffer, C. M., & Borduin, C. M. (2005). Long-term follow-up to a randomized clinical trial of multisystemic therapy with serious and violent juvenile offenders. *Journal of Consulting and Clinical Psychology*, *73*(3), 445-453. doi:2005-06517-007

Schoenwald, S. K., & Henggeler, S. W. (2004). A public health perspective on the transport of evidence-based practices. *Clinical Psychology: Science and Practice*, *11*(4), 360-363.
doi:10.1207/s15374424jccp3301_10

Schawo, S. J., van Eeren, H., Soeteman, D. I., van der Veldt, M.-C., Noom, M. J., Brouwer, W. B. F., . . . Hakkaart, L. (2012). Framework for modelling the cost-effectiveness of systemic interventions aimed to reduce youth delinquency. *Journal of Mental Health Policy and Economics*, *15*(4), 187-196.

Scorecard (n.d.). The pollution information site: Compare your county. Retrieved from
<http://scorecard.goodguide.com/community/cm-compare.tcl?fips1=06013&zip1=94509&zip2=94556&fips2=06013>

Sexton, T. L. & Alexander, J. F. (2004). *Functional Family Therapy clinical training manual*. Annie E. Casey Foundation.

Sexton, T.L. & Turner, C. W. (2011). The effectiveness of functional family therapy for youth with behavioral problems in a community practice setting. *Couple and Family Psychology: Research and Practice*, *1*(S), 3-15. doi: 10.1037/2160-4096.1.s.3

Shade, K. & Ceric, D. (2016). *Contra Costa County Probation's 'Thinking for a Change' Program Evaluation*. Evidence Based Practice to Improve Public Safety (EBP-TIPS) grant, awarded to Contra Costa County Probation, 2013-2015

Shade, K. & Ceric, D. (2017). *Girls in Motion (GIM) Treatment Program, Final Evaluation Report, 2017*. Internal Report for Contra Costa County Probation

Sheidow, A. J., McCart, M. R., & Davis, M. (2016). Multisystemic Therapy for emerging adults with serious mental illness and justice involvement. *Cognitive and Behavioral Practice, 23*(3), 356–367. <http://doi.org/10.1016/j.cbpra.2015.09.003>

Sowers, W., George, C., & Thompson, K. (1999). Level of care utilization system for psychiatric and addiction services (LOCUS): a preliminary assessment of reliability and validity. *Community Mental Health Journal, 35*(6), 545-563. <https://doi.org/10.1023/A:1018767403107>

Stamm, B.H. (2010). *The Concise ProQOL Manual*. Pocatello, ID: ProQOL.org

State of California, Health & Human Services Department (2016, Sept 30). *All-county letter no: 16-85*. Retrieved from <http://www.cdss.ca.gov/lettersnotices/EntRes/getinfo/acl/2016/16-85.pdf>

State of California (2018). *Welfare and Institutions Code, Article 20, section 777*. Retrieved from http://leginfo.legislature.ca.gov/faces/codes_displaySection.xhtml?sectionNum=777.&lawCode=WIC

Stoltzfus, J.C. (2011). Logistic regression: A brief primer. *Annals of Emergency Medicine, 18*(10), 1099-1104. <https://doi.org/10.1111/j.1553-2712.2011.01185.x>

Thornberry, T. P., Freeman-Gallant, A., & Lovegrove, P. J. (2009). The impact of parental stressors on the intergenerational transmission of antisocial behavior. *Journal of Youth and Adolescence, 38*(3), 312-322. doi: 10.1007/s10964-008-9337-0

Timmons-Mitchell, J., Bender, M. B., Kishna, M. A., & Mitchell, C. C. (2006). An independent effectiveness trial of multisystemic therapy with juvenile justice youth. *Journal of Clinical Child and*

Adolescent Psychology, 35(2), 227-236. doi:10.1207/s15374424jccp3502_6

Trulson, C. R., Marquart, J. W., Mullings, J. L., & Caeti, T. J. (2005). In between adolescence and adulthood: Recidivism outcomes of a cohort of state delinquents. *Youth Violence and Juvenile Justice*, 3(4), 355-387. doi 10.1177/1541204005278802

University of Cincinnati Corrections Institute (n.d.) Core Correctional Practices (CCR) training overview. Retrieved from https://cech.uc.edu/centers/ucci/services/trainings/changing_offender_behavior/ccptrainingoverview.html

Unruh, D.K., Gau, J.M. & Waintrup, M.G. (2009). An exploration of factors reducing recidivism rates of formerly incarcerated youth with disabilities participating in a re-entry intervention. *Journal of Child and Family Studies*, 18, 284. <https://doi.org/10.1007/s10826-008-9228-8>

van Batenburg-Eddes, T., Butte, D., van de Looij-Jansen, P., Schiethart, W., Raat, H., de Waart, F., & Jansen, W. (2012). Measuring juvenile delinquency: How do self-reports compare with official police statistics? *European Journal of Criminology*, 9(1), 23-37. <https://doi.org/10.1177/1477370811421644>

van der Ende, J., & Verhulst, F. C. (2005). Informant, gender and age differences in ratings of adolescent problem behaviour. *European Child and Adolescent Psychiatry*, 14(3), 117-126. doi 10.1007/s00787-005-0438-y

van der Put, C. E., Deković, M., Hoeve, M., Stams, G. M., van der Laan, P. H., & Langewouters, F. M. (2014). Risk assessment of girls: Are there any sex differences in risk factors for re-offending and in risk profiles? *Crime and Delinquency*, 60(7), 1033-1056. doi 10.1177/0011128710384776

Viljoen, J. L., Shaffer, C. S., Gray, A. L., & Douglas, K. S. (2017). Are adolescent risk assessment tools sensitive to change? A framework and examination of the SAVRY and the YLS/CMI. *Law and Human Behavior*, 41(3), 244-257. doi:10.1037/lhb0000238

Waldron, H. B., Slesnick, N., Brody, J. L., Turner, C. W. & Peterson, T. R. (2001). Treatment outcomes for adolescent substance abuse at 4- and 7-month assessments. *Journal of Consulting and Clinical Psychology*, 69(5), 802-813. doi: 10.1037/0022-006X.69.5.802

Wells, M. G., Burlingame, G. M., Lambert, M. J. & Hoag, M. J. (1996). Conceptualization and measurement of patient change during psychotherapy: Development of the Outcome Questionnaire and Youth Outcome Questionnaire. *Psychotherapy*, 33, 275-283. <https://doi.org/10.1037/0033-3204.33.2.275>

Wen, L., Badgett, R., & Cornell, J. (2005). Number needed to treat: A descriptor for weighing therapeutic options. *American Journal of Health System Pharmacy*, 62(19), 2031-2036. doi:10.2146/ajhp040558

Westfall, P., & Henning, K. (2013). *Understanding advanced statistical methods*. Boca Raton, FL: CRC Press.

Wilbur, S. (2016). *Final CA-LOCUS training*. [PowerPoint Slides].

Wiltsey-Stirman, S., Kimberly, J., Cook, N., Calloway, A., Castro, F., & Charns, M. (2012). The sustainability of new programs and innovations: A review of the empirical literature and recommendations for future research. *Implementation Science* 7, 17. <http://doi.org/10.1186/1748-5908-7-17>

Wolfgang, M., Figlio, R. & Sellin, T. (1972). *Delinquency in a birth cohort*. Chicago, IL: University of Chicago Press.

Youngstrom, E. A. (2014). A primer on receiver operating characteristic analysis and diagnostic efficiency statistics for pediatric psychology: We are ready to ROC. *Journal of Pediatric Psychology*, 39(2), 204–221. <http://doi.org/10.1093/jpepsy/jst062>

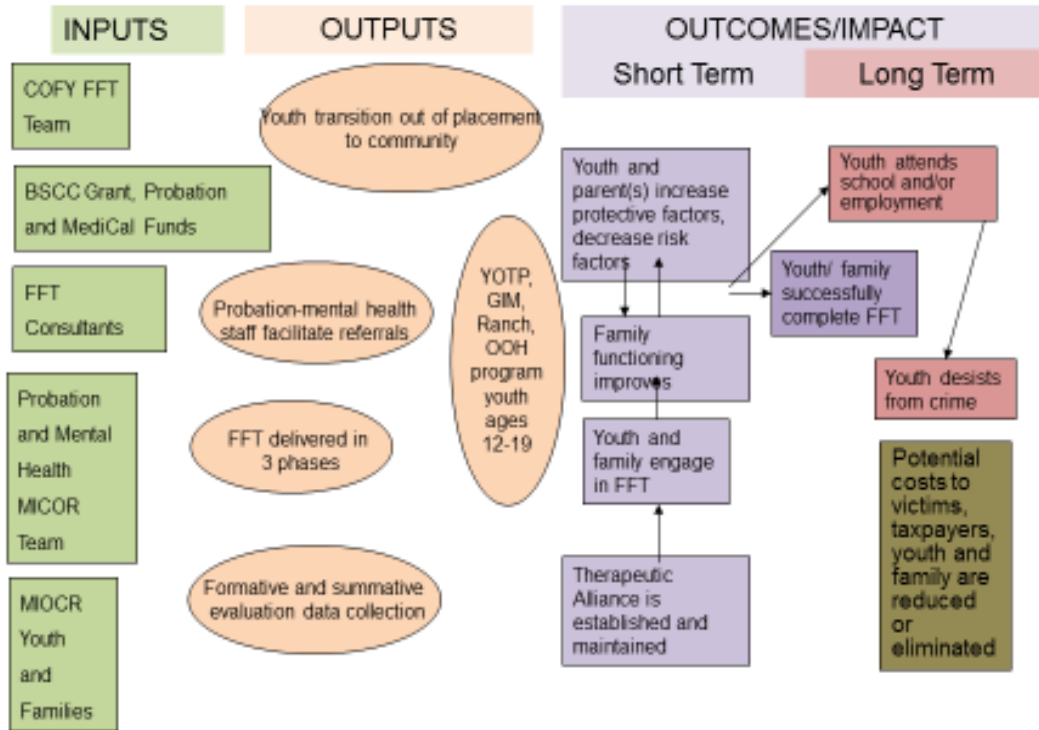
Zazzali, J. L., Sherbourne, C., Hoagwood, K. E., Greene, D., Bigley, M. F., & Sexton, T. L. (2008). The adoption and implementation of an evidence based practice in child and family mental health

services organizations: A pilot study of Functional Family Therapy in New York State. *Administration and Policy in Mental Health and Mental Health Services Research*, 35(1), 38-49. doi 10.1007/s10488-007-0145-8

APPENDIX A

Logic Model for Contra Costa County's MIOCR-funded TOSO Project

Logic Model for MIOCR Project



APPENDIX B

Child and Adolescent Level of Care Utilization System (CALOCUS) Worksheet

CALOCUS WORKSHEET

Rater Name _____ Date _____

Please check the applicable ratings within each dimension and record the score in the lower right hand corner. Total your score and determine the recommended level of care using either the Placement Grid or the Decision Tree.

<p>I. Risk of Harm</p> <p><input type="checkbox"/> 1. Low Risk of Harm</p> <p><input type="checkbox"/> 2. Some Risk of Harm</p> <p><input type="checkbox"/> 3. Significant Risk of Harm</p> <p><input type="checkbox"/> 4. Serious Risk of Harm</p> <p><input type="checkbox"/> 5. Extreme Risk of Harm</p> <p style="text-align: right;">Score _____</p>	<p>IV-B. Recovery Environment - Environmental Support</p> <p><input type="checkbox"/> 1. Highly Supportive Environment</p> <p><input type="checkbox"/> 2. Supportive Environment</p> <p><input type="checkbox"/> 3. Limited Support in Environment</p> <p><input type="checkbox"/> 4. Minimally Supportive Environment</p> <p><input type="checkbox"/> 5. No Support in Environment</p> <p style="text-align: right;">Score _____</p>
<p>II. Functional Status</p> <p><input type="checkbox"/> 1. Minimal Functional Impairment</p> <p><input type="checkbox"/> 2. Mild Functional Impairment</p> <p><input type="checkbox"/> 3. Moderate Functional Impairment</p> <p><input type="checkbox"/> 4. Serious Functional Impairment</p> <p><input type="checkbox"/> 5. Severe Functional Impairment</p> <p style="text-align: right;">Score _____</p>	<p>V. Resiliency and Treatment History</p> <p><input type="checkbox"/> 1. Full Resiliency and/or Response to Treatment</p> <p><input type="checkbox"/> 2. Significant Resiliency and/or Response to Treatment</p> <p><input type="checkbox"/> 3. Moderate or Equivocal Resiliency and/or Response to Treatment</p> <p><input type="checkbox"/> 4. Poor Resiliency and/or Response to Treatment</p> <p><input type="checkbox"/> 5. Negligible Resiliency and/or Response to Treatment</p> <p style="text-align: right;">Score _____</p>
<p>III. Co-Morbidity</p> <p><input type="checkbox"/> 1. No Co-Morbidity</p> <p><input type="checkbox"/> 2. Minor Co-Morbidity</p> <p><input type="checkbox"/> 3. Significant Co-Morbidity</p> <p><input type="checkbox"/> 4. Major Co-Morbidity</p> <p><input type="checkbox"/> 5. Severe Co-Morbidity</p> <p style="text-align: right;">Score _____</p>	<p>VI-A. Acceptance and Engagement - Child/Adolescent</p> <p><input type="checkbox"/> 1. Optimal</p> <p><input type="checkbox"/> 2. Constructive</p> <p><input type="checkbox"/> 3. Obstructive</p> <p><input type="checkbox"/> 4. Adversarial</p> <p><input type="checkbox"/> 5. Inaccessible</p> <p style="text-align: right;">Score _____</p>
<p>IV-A. Recovery Environment - Environmental Stress</p> <p><input type="checkbox"/> 1. Minimally Stressful Environment</p> <p><input type="checkbox"/> 2. Mildly Stressful Environment</p> <p><input type="checkbox"/> 3. Moderately Stressful Environment</p> <p><input type="checkbox"/> 4. Highly Stressful Environment</p> <p><input type="checkbox"/> 5. Extremely Stressful Environment</p> <p style="text-align: right;">Score _____</p>	<p>VI-B. Acceptance and Engagement - Parent/Primary Caretaker</p> <p><input type="checkbox"/> 1. Optimal</p> <p><input type="checkbox"/> 2. Constructive</p> <p><input type="checkbox"/> 3. Obstructive</p> <p><input type="checkbox"/> 4. Adversarial</p> <p><input type="checkbox"/> 5. Inaccessible</p> <p style="text-align: right;">Score _____</p>
<p>Composite Score <input style="width: 50px;" type="text"/></p>	<p>Level of Care Recommendation <input style="width: 50px;" type="text"/></p>

APPENDIX C

Categories for Level of Care based on Composite CALOCUS/LOCUS Score

Level of Care	Description	Composite CALOCUS/ LOCUS Score
Zero	Basic services for prevention and maintenance	7-9
One	Recovery maintenance and health management	10-13
Two	Low-intensity community based services	14-16
Three	High-intensity community based services	17-19
Four	Medically monitored community based services	20-22
Five	Medically monitored residence based services	23-27
Six	Medically managed residence based services	28+

APPENDIX D

Contra Costa County Residence Zip Codes for TOSO Youth Referrals and those with Sustained Petitions

